



Master's thesis

Solid Earth Geophysics

A New Ground Motion Prediction Equation for Otaniemi, Espoo, South-Finland

Ahti Ilmari Voutilainen

2021

Supervisors:

M.Sc. Tommi Vuorinen

Ph.D. Päivi Mäntyniemi

Master's Programme in Geology and Geophysics

Faculty of Science



Koulutusohjelmasta vastaava tiedekunta – Fakulteten ansvarig för programmet – Faculty responsible for the programme		Tutkinnon myöntävä tiedekunta – Fakulteten som beviljar examen – Faculty granting the degree	
Faculty of Science		Faculty of Science	
Tekijä – Författare – Author Ahti Ilmari Voutilainen			
Tutkielman otsikko – Avhandlingens titel – Title of thesis A New Ground Motion Prediction Equation for Otaniemi, Espoo, South-Finland			
Koulutusohjelma – Utbildningsprogram – Study programme Master's Programme of Geology and Geophysics			
Tutkielman taso – Avhandlingens nivå – Level of the thesis Master's thesis	Aika – Datum – Date 29.04.2021	Sivumäärä – Sidoantal – Number of pages 55	
Tiivistelmä – Referat – Abstract <p>A new ground motion prediction equation, named ON21, is solved for the ST1 Deep Heat enhanced geothermal system in Otaniemi, city of Espoo, Finland. The raw data from seismic events, that occurred during the stimulation of 2018, is processed, instrument response is removed, and frequency domain is used to obtain peak ground displacement, velocity, and acceleration.</p> <p>A database with 20,768 ground motion recordings from 204 events is compiled and used to solve a ground motion prediction equation for peak ground velocity and acceleration for vertical and horizontal movement. The model has a magnitude range from 0.0 to 1.8 on the scale of local magnitude used in Finland, and hypocentral distances of 0 km to 20 km. A 1σ value of 0.60 for vertical peak ground velocity model is lower than the 1σ of the models previously in use at Otaniemi and its surrounding areas.</p> <p>It is observed that the azimuth between the strike of the fault causing the earthquakes and the station recording the events seems to affect the peak ground motion values at a hypocentral distance of no more than 10 km, and beyond that the magnitude and distance are the dominant factors in the peak ground motion values. The new ground motion prediction equation model ON21 should be tested with ground motion data from earthquakes that occurred during the 2020 stimulations to assess its usefulness in predicting peak ground motion values, and to further study the effect of azimuth on the peak ground motion values.</p>			
Avainsanat – Nyckelord – Keywords ground motion prediction equation, peak ground motion value, Otaniemi, seismic event, earthquakes			
Säilytyspaikka – Förvaringställe – Where deposited HELDA – Digital repository for University of Helsinki			
Muita tietoja – Övriga uppgifter – Additional information 26 figures, 4 tables			



Tiedekunta – Fakultet – Faculty Matemaattis-luonnontieteellinen tiedekunta		Koulutusohjelma – Utbildningsprogram – Degree programme Geologian ja geofysiikan maisteriohjelma	
Opintosuunta – Studierikting – Study track solid earth geophysics–kiinteän maan geofysiikka			
Tekijä – Författare – Author Ahti Ilmari Voutilainen			
Työn nimi – Arbetets titel – Title A New Ground Motion Prediction Equation for Otaniemi, Espoo, South-Finland			
Työn laji – Arbetets art – Level Pro-gradu	Aika – Datum – Month and year 29.04.2021	Sivumäärä – Sidoantal – Number of pages 55	
Tiivistelmä – Referat – Abstract <p>Uusi maanliikkeen maksimiarvoja ennustava malli, ON21, ratkaistaan ST1 Deep Heat:n geolämpövoimalalle Otaniemeen, Espoon kaupunkiin, Etelä-Suomeen. Raakadataa maanjäristyksistä, jotka sattuivat 2018 stimuloinnin yhteydessä, käsitellään, instrumentin vaste poistetaan ja taajuusaluetta käytetään integrointiin ja derivointiin, jotta saadaan ratkaistuksi maanliikkeen maksimisiirtymät, -nopeudet ja -kiihtyvyydet.</p> <p>CSV-tiedostoon kerätään maksimi maanliikearvot 20 768 nauhoitteesta, jotka ovat 204 maanjäristyksestä, jotka sattuivat 2018 stimuloinnin yhteydessä. Näitä arvoja käytetään maanliikkeen ennustavan mallin ratkaisuksi maksiminopeuden ja -kiihtyvyyden arvoille vertikaali- ja horisontaalisuunnissa. Malli toimii Suomessa käytetyn paikallisen magnitudin välillä 0,0–1,8 ja hyposentrin etäisyyksillä 0–20 km. Maksimi vertikaalinopeuden mallin keskihajonta 0,60 on pienempi, kuin aikaisemmin Otaniemen lämpövoimalalla käytettyjen mallien keskihajonnat.</p> <p>Atsimuutti järistyksiä aiheuttavan rakovyöhykkeen kulun ja maanliikettä nauhoittavan aseman välillä näyttää vaikuttavan maanliikkeen maksimiarvoihin alle 10 km:n etäisyyksillä hyposentristä. Yli 10 km etäisyyksillä etäisyyden ja magnitudin vaikutus peittää alleen atsimuutin vaikutuksen. Uusi maanliikkeen maksimiarvoja ennustava malli ON21 tulisi testata vuonna 2020 stimuloinnin aiheuttamien maanjäristysten aiheuttamilla maanliikkeen maksimiarvoilla, jotta sen hyödyllisyyttä ennustavana mallina voidaan arvioida, ja jotta atsimuutin vaikutusta voidaan tutkia tarkemmin.</p>			
Avainsanat – Nyckelord – Keywords maanliikkeen ennustava malli, maanliikkeen maksimiarvo, Otaniemi, maanjäristys			
Säilytyspaikka – Förvaringställe – Where deposited HELDA – Digital Repository for University of Helsinki			
Muita tietoja – Övriga uppgifter – Additional information 26 kuvaa, 4 taulukkoa			



1. INTRODUCTION	4
1.1 Study Area: A Brief Geological Overview of Otaniemi and Its Surrounding Areas	5
2. THEORETICAL BACKGROUND	8
2.1 Ground Motion Prediction Equations	8
2.2 Forms of Ground Motion Prediction Equations–What Factors into Ground Motion?	9
2.3 Instrument Response and Its Removal.....	12
2.4 Fourier Transform and Frequency Domain	14
2.5 Ground Motion–From Physical Motion to a Digital Signal	17
2.6 Ground Motion–From Digital Signal to Useful Data.....	18
3. DATASET	19
4. METHODS–FROM RAW DATA TO A MODEL.....	22
5. RESULTS	28
6. DISCUSSION.....	44
6.1 The Peak Ground Motion Database.....	44
6.2 Model ON21 Ground Motion Prediction Equation.....	45
6.3 Effect of Azimuth	48
7. CONCLUSIONS.....	49
8. ACKNOWLEDGEMENTS	51
9. REFERENCES	52
APPENDIX 1: PYTHON FUNCTIONS.....	56
APPENDIX 2: EXAMPLE OF ST1-18_DATABASE_PGM.CSV	61
APPENDIX 3: VERTICAL PEAK GROUND MOTION VARIATIONS BY MAGNITUDE FROM ST1-18_DATABASE_PGM.CSV	63

1. INTRODUCTION

In this master's thesis, a new Ground Motion Prediction Equation (GMPE) is solved to supplement existing methods of probabilistic seismic hazard analysis (PSHA) for the ST1 Deep Heat Enhanced Geothermal System (EGS) at Otaniemi, city of Espoo, Finland. Arup (2018) propose solving a more accurate GMPE that uses high-quality surface data from the Institute of Seismology of the University of Helsinki (ISUH). The new model is used to calibrate an existing traffic light system (TLS) that is used to monitor the seismicity caused by the EGS. A more accurate model of peak ground motion and attenuation of seismic energy would allow the operation of less stations to be used at carefully selected locations, instead of a many large arrays at different locales.

With the push for more clean energy sources, geothermal energy is rising as an alternative source for heating to other energy sources (Buijze et al. 2019, Veikkolainen et al. 2021). Drilling and pumping water deep into bedrock increases a seismic risk (Veikkolainen et al. 2021), requiring a monitoring of the seismicity during the building and operation of any system of hydraulic fracturing or geothermal stimulation (Schultz et al. 2020); there are geothermal well projects that have been shut down completely due to an increase in seismicity related to the operating of an EGS (e.g. in Switzerland; Giardini 2009). A good monitoring system with accurate predictive models can help mitigate these risks (Ader et al. 2020, Schultz et al. 2020). However, historically the capital region of Finland (CRF), where the city of Espoo is, has been a seismically quiescent area and solving for an accurate GMPE has been challenging due to a lack of events to draw a database from. Previously GMPEs in Finland have been solved for either the entire Fennoscandia (Vuorinen et al. 2018) or on areas with nuclear power plants (Fülöp et al. 2019). As GMPEs are regression models of ground motion, they require prior earthquake data (empirical or theoretical) to which a model is fitted (Campbell 2003). With the induced earthquakes at Otaniemi, a demand to solve for an accurate GMPE has risen as well as events to base the model on.

While some models draw from attenuation models made for seismologically similar areas (e.g. Pezeshk et al. 2011, Vuorinen et al. 2018), at low magnitudes and relatively low peak ground motion values (in Otaniemi's case M_L [local magnitude] < 2 , PGV [Peak Ground Velocity] $< 3 \text{ mms}^{-1}$; Ader et al. 2020), it is important to emphasize the target site's ground motion recordings (Campbell 2010, Douglas 2011). Campbell (2010) concludes that at small magnitudes ($M < 5.5$), without correct site-specific observations, GMPEs tend to predict too large ground motion at great distances. In the case of this study, because the focus is on a 20 km radius, this is unlikely to become an issue. While PSHAs most commonly use peak ground acceleration (PGA; Lekshmy and Raghukanth 2021), in this study a model for vertical and horizontal PGV and PGA are solved

This study has three main objectives:

1. To collect a database of peak ground motion values using earthquake data from ST1 Deep Heat EGS's 2018 stimulation
2. To solve a new GMPE model, named ON21, for vertical and horizontal PGV and PGA to be used at Otaniemi and its surrounding areas and
3. To study if the azimuth between the strike of the earthquake fault and the recording station affects the peak ground motion values.

The database collection is done using Python programs with ObsPy-module, the GMPE's constants are solved by fitting Campbell's (2003) model of GMPE into the peak ground motion data, and the effect of azimuth is studied by comparing the peak ground motion data with the magnitudes of the events, the hypocentral distances of the recordings, and azimuths of the recordings.

1.1 Study Area: A Brief Geological Overview of Otaniemi and Its Surrounding Areas

The bedrock of Southern Finland is characterised by high temperature, low pressure metamorphism. Metamorphism is mostly amphibolite facies (4–5 kbar, 650–700 °C) at

late Svecofennian time (1.85–1.78 Ga). It is likely the result of crustal thickening, heating, and collapse (Kukkonen and Lauri 2009). This has left the bedrock with mostly crystallised material: granitoids, late-orogenic granites and paragneisses dominate the area around the EGS at Otaniemi and Laajalahti (Kukkonen and Lauri 2009). Veikkolainen and Kukkonen (2019) give interpolated estimates for average heat production and density at the capital region of Finland. The heat production is interpolated from U-Th-K-data and is between 1.2 and 1.4 μWm^{-3} , and density is 2800–2925 kgm^{-3} with relatively high standard deviations due to changes in bedrock at any given place. The Mohorovičić discontinuity depth is given by Kukkonen et al. (2008) and is roughly 50 km at the CRF, which is typical for a craton.

Figure 1 shows the main lithological units at the Otaniemi borehole (Geological Survey of Finland 2017). The bedrock around the Otaniemi borehole and towards south-west is microcline granite, with amphibolites and biotite paragneisses surrounding to all other directions. North of Laajalahti, under which the borehole continues and where most of the seismic events occurred, are more metamorphic rocks: quartz feldspar paragneisses cut a narrow path with more amphibolites further north, and at 5 km distance to the north is granodiorite. Most of the capital region is mixed plutonic rocks and gneisses, with the granitoids making up most of the charted bedrock. Figure 1 also shows the main fault-lines as violet lines. The faults show some lineation in a north-east–south-west direction close to the Otaniemi borehole, as do the quartz feldspar paragneiss (yellow) and its surrounding units. However, Leonhardt et al. (2021) conclude that the strike of the fault causing the earthquakes at Otaniemi is north-north-west–south-south-east oriented. Hillers et al. (2020) had previously found that the horizontal S-waves radiated the strongest towards north-east and north-west, whereas the vertical P-waves were the strongest directly north, east and west of the borehole. Comparing the findings with the lithological map, it is unlikely that the different rock-types have a significant impact to the relative peak ground motion values around Otaniemi. Additionally Campbell (2010) gives shear-wave velocities for hard rock sites at $V_S = 2000 \text{ ms}^{-1}$, and Schön (2011) gives shear-wave velocities for granite ($V_S = 2600\text{--}3200 \text{ ms}^{-1}$), gneisses ($V_S = 2800\text{--}2900 \text{ ms}^{-1}$) and diorite ($V_S = 2900\text{--}3200 \text{ ms}^{-1}$). Thus, a hard-rock environment is assumed for Otaniemi and its surrounding areas.

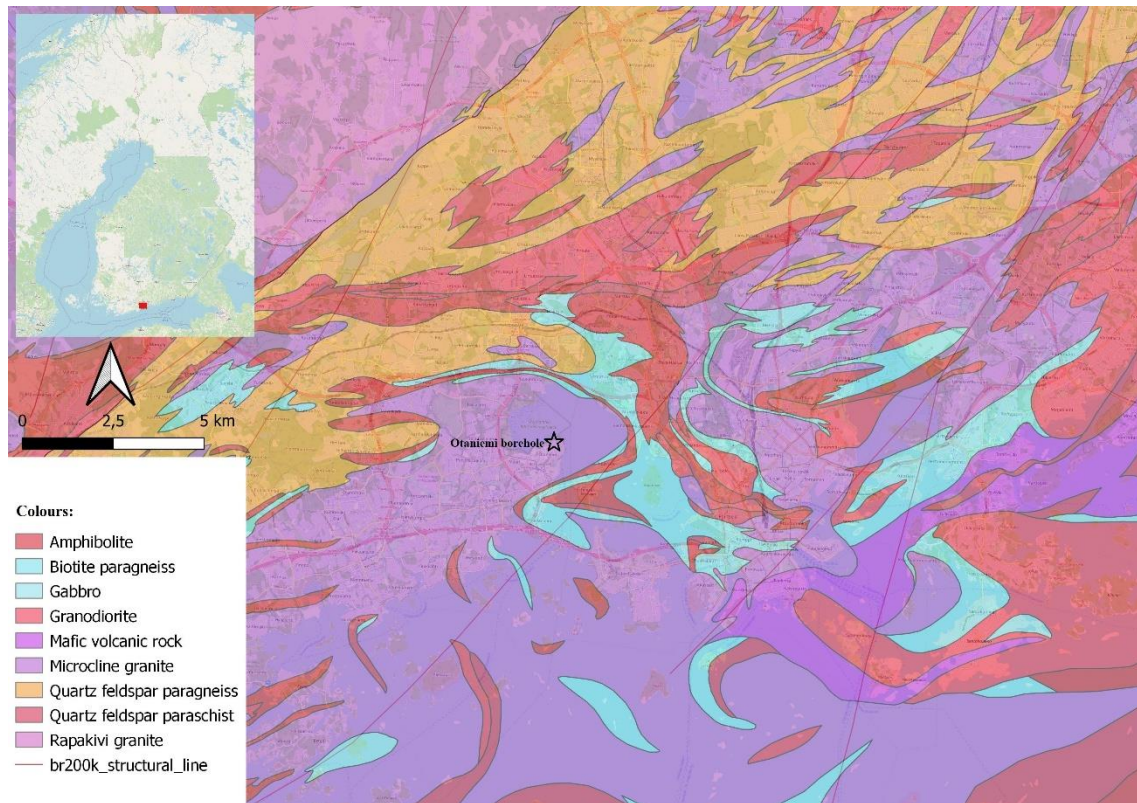


Figure 1. The main lithological units at the CRF (bedrock map by Geological Survey of Finland 2017).

The tectonic environment affects the GMPEs in different ways (Campbell 2003, Kawase 2003). Campbell (2003) gives four categories for tectonic environments, of which category two, shallow crustal earthquakes in stable continental regions, describes best the induced earthquakes of Otaniemi. Site location plays a role as well (Kawase 2003). GMPEs are, by design, estimations of “strong ground motion on level ground in free field” (Campbell 2003). This mainly puts constraints on how the data used in forming a GMPE should be recorded but not necessarily the values of the parameters. The recording equipment should not be on large structures, sites with varying topography or below ground. Often such recordings are not included in the creation of a GMPE, with some exceptions e.g. due to the solving of the models requiring large databases (Campbell 2003, Kawase 2003).

2. THEORETICAL BACKGROUND

The next chapter discusses the theoretical insights required to understand and solve for functioning and reliable ground motion prediction equations. GMPEs are discussed, what factors into ground motion, what is the specific model solved in this study and how ground motion data is produced.

2.1 Ground Motion Prediction Equations

GMPEs or attenuation relations are mathematical approximations of stochastic ground motion caused by a seismic event (Anderson 2003, Campbell 2003). Ground motion is often described with peak parameters (PGV: Peak Ground Velocity, PGA: Peak Ground Acceleration, among others; Anderson 2003, Campbell 2003) that are derived from observations of ground motion either on the site of the specific GMPE or on a tectonically similar site (such as Fennoscandia and Eastern North America) and in some cases theoretical data (Campbell 2003, Pezeshk et al. 2011, Vuorinen et al. 2018), making any GMPE a regression fit (Campbell 2003). Empirical GMPEs assume that, at any given location, if the mechanisms of two earthquakes and their magnitudes are similar to one another, they cause similar ground movement as well (Douglas and Aochi 2008). In practice, a GMPE shows the peak ground motions' mean expected value in relation to the parameters provided in its equation, with the assumption that the ground motions from multiple events are distributed in a Gaussian manner with its randomness caused by aleatory variability, i.e. the randomness of distribution of the peak ground motion (Bolt and Abrahamson 2003, Campbell 2003). Some models also consider the epistemic uncertainties, or the known uncertainties of the models (Bolt and Abrahamson 2003).

GMPEs have been used as tools for PSHA in areas of both natural and induced seismicity, especially if the risk caused by seismicity is somehow increased (due to a presence of a nuclear power plant, urban environment, high population density, etc.; e.g. Douglas and Aochi 2008, Fülöp et al. 2019, Lekshmy and Raghukanth 2021). Similarly mining and drilling projects conducted in the vicinity of urban areas may be a cause for increased seismic hazard, prompting the use of a GMPE as a tool for predicting the likely peak

ground motions caused by different events (Ader et al. 2020, Schultz et al. 2020). For instance, Arup (2018) propose the design of a GMPE to calibrate a TLS created for monitoring the induced seismicity at the EGS of ST1 in Otaniemi, city of Espoo, Finland. A well calibrated system would allow easier operation of an EGS and provide better estimates of maximum ground movement, which are both important for PSHA, local preparation and the public acceptance of any project likely to induce seismic events (Kwiatek et al. 2019, Ader et al. 2020).

2.2 Forms of Ground Motion Prediction Equations—What Factors into Ground Motion?

For areas with a low mean magnitude and relatively short hypocentral distances, site-specific ground motion is observed (Douglas 2011). After considering site conditions and local magnitude scaling, weak ground motion seems to attenuate differently in relation to magnitude and hypocentral distance than strong ground motion (Bommer et al. 2007, Douglas 2011). GMPEs are characterized by source and strong-motion parameters (Anderson 2003, Campbell 2003). Source parameters are magnitude of the event, type of faulting and focal mechanism, stress drop, and source directivity and radiation pattern (Campbell 2003). Strong-motion parameters can be expressed as either peak time-domain parameters or peak frequency-domain parameters. Peak time-domain parameters include the aforementioned PGV and PGA as well as peak ground displacement (PGD), the three being derivatives of each other. Peak frequency-domain parameters include pseudoacceleration (PSA) and pseudovelocity (PSV), with the two being related to each other by a relative displacement expression (Campbell 2003).

The GMPE solved in this study is based on Campbell's (2003) method, and the model given by them. The method uses a predetermined attenuation relation function which is fitted into a database using the least squares method to solve for the unknown parameters. The least squares method is a method of regressive analysis, that is used when there are more equations than unknowns. In situations like these, the equations are often inconsistent in their unknowns. The model is fitted into the data so that the residual squares (squares of the difference between model and a point of data) are as small as

possible (Campbell 1981, Campbell 2003). The ground motion database will be used to modify the parameters of this model, and as such the model will estimate the ground motion at distances and magnitudes found from the database. However, due to its empirical nature, the GMPE cannot predict ground motion properly at distances or magnitudes beyond the database with uncertainty increasing towards the limits of the model (Bommer et al. 2007).

Outliers in the database can also have an impact on the model (Hansen et al. 2013). Due to the least squares method, a data-bias is present, making the model fit best in areas of a lot of data points (in this case recordings from seismic events) unless the model is weighted (Hansen et al. 2013, Vuorinen et al. 2018). In this study such areas may be the distances from the events that have the most stations, and the events of certain magnitudes that have the most recordings. During the analysis of the 2018 data, the events of $M_L = 0.5$ and above were prioritized, meaning that not many events of $M_L = 0.0$ or close were analysed in comparison. Contrasting with stimulation of 2020, most events had negative local magnitude, which means the dataset would have a bias towards weaker events.

Besides Campbell's (2003) model, other models for GMPE exist, as well as other methods for creating the models. Atkinson's (2008) method (the "referenced empirical approach") suggests using an existing model from a seismologically similar area and multiplying its constants with parameters received from using a database specific to the new model's site. This method has the benefit of using the databases of both the old GMPE's site and the new site (Atkinson 2008). Hybrid methods, combining the two approaches, exist as well (Douglas and Aochi 2008).

Campbell (2003) expresses GMPEs with either a logarithmic form or as a function of Y :

$$\ln Y = c_1 + c_2 M - c_3 \ln R - c_4 r + c_5 F + c_6 S + \varepsilon \quad (eq\ 1)$$

or in exponential form

$$Y = c_1 e^{c_2 M} R^{-c_3} e^{-c_4 r} e^{c_5 F} e^{c_6 S} e^{\varepsilon} \quad (eq\ 2),$$

where Y is the strong-motion parameter of interest (e.g. PGD, PGV or PGA; Kafaei Mohammadnejad et al. 2012), M is the event magnitude (this study uses local magnitude scale in use in Finland), r is distance between source and receiver, F is a parameter dependent on the type of faulting causing the event, or the azimuth to the plane of rupture, S is a parameter dependent on the local site conditions (including the bedrock and soil conditions), R is a magnitude-related distance term, and ε is the random error of estimate, or the standard deviation of $\ln Y$ (Campbell 2003).

Due to the computational difficulties and epistemic uncertainties in trying to solve for multiple coefficients, a few simplifying assumptions are made. Because the induced events at Otaniemi have relatively low mean magnitude, the term r is used instead of R . Weak earthquakes can be approximated with having a point-source, which is why either r_{epi} or r_{hypo} are used for epicentre and hypocentre respectively. This study will use the hypocentral distance r_{hypo} , hence referred to as r . The R -term is more useful when dealing with strong earthquakes with large ruptures that are difficult to assume having a point-source (Campbell 2003).

The F denoting the faulting and azimuth is omitted as well, since there were very few events that had a different faulting mechanism (Veikkolainen et al. 2020). However, the possible effect of the azimuth is discussed briefly, while it is not implemented into the model. Previous studies indicate that at distances less than roughly 20 km, there are azimuthal differences in the energy of the ground motion (e.g. Hillers et al. 2020).

Site conditions S is also omitted from the solution because the area around Otaniemi can be considered a ‘hard-rock’ environment. Base 10 logarithm is used instead of the base e logarithm for compliance with previous work (e.g. Vuorinen et al. 2018). Local magnitude M_L is used because the solutions for the events have their magnitude as local magnitude instead of moment magnitude.

With these assumptions the final equation is:

$$\log_{10}(PGV) = c_1 + c_2 M_L - c_3 r + \varepsilon \text{ (eq 3)},$$

where the regression coefficients c_{1-3} are solved in this study. C_1 is the adjustment constant, c_2 is the magnitude's multiplication constant and c_3 is the distance's multiplication constant. The error factor ε is the standard error of $\log_{10}(PGV)$ which is assumed to follow a Gaussian distribution:

$$\begin{aligned} \varepsilon = \sigma_{\ln Y} &= \sqrt{1/(n-p) \sum (\ln Y_i - \overline{\ln Y_i})^2} \\ &= \sqrt{1/(n-3) \sum (\log_{10}(PGV_i) - \overline{\log_{10}(PGV_i)})^2} \text{ (eq 4)}, \end{aligned}$$

where n is the number of observations (recordings in this case), p is the number of regression coefficients (3 in this case), $\ln Y_i$ or $\log_{10}(PGV)$ is the i^{th} event and $\overline{\ln Y_i}$ or $\overline{\log_{10}(PGV)_i}$ is the predicted value of the same event, i.e. the value the model gives (Campbell 2003). Error values for magnitude and distance are received from the analysis programs and are already in the database. From here on, the error will be marked with σ instead of ε .

2.3 Instrument Response and Its Removal

Due to slowness of mass, traction, and the stiffness of the spring, the unprocessed data of a seismogram does not depict ground movement. Instrument response refers to the part of the seismogram's data that is not ground movement but caused by the instrument itself (Asch 2009). In mathematical terms a seismogram's data $T(j\omega)$ can be expressed as the convolution $*$ of ground movement $F(j\omega)$ and instrument response $H(j\omega)$. In this case, the operating is in frequency domain (Gadallah and Fisher 2009). By Fourier transform the signal can be changed from time to frequency domain, where the convolution operation transforms into a multiplication. Deconvolving is now done by dividing the ground motion with the instrument response:

$$T(j\omega) = F(j\omega) * H(j\omega) \rightarrow F(j\omega) = \frac{T(j\omega)}{H(j\omega)} \text{ (eq 5)},$$

where the $j\omega$ is the frequency domain of the signal (Scherbaum 2001, Wielandt 2012). If the system response that is convolving the earthquake signal can be mathematically

expressed, the expression can be used to divide, or deconvolve, the measured signal at every time step of the discrete signal, leaving only the ground movement.

In a simple case of one spring system, Scherbaum (2001) considers instrument response to have three components: the force trying to prevent the change of acceleration (Equation 6.1), the stiffness of the spring as described by Hooke's law (Equation 6.2), and some damping effect (Equation 6.3). Calculating the three together in equilibrium provides Equation 6.4 that solves into the response. Equations 6.1–6.4 are from Scherbaum (2001).

$$F_h = -m\ddot{u}_m(t) \text{ (eq 6.1, force resisting change of acceleration)}$$

$$F_j = -kx_r(t) \text{ (eq 6.2, force of the spring)}$$

$$F_v = -D\dot{x}_m(t) \text{ (eq 6.3, force causing dampening)}$$

$$\rightarrow -m\ddot{u}_m(t) - kx_r(t) - D\dot{x}_m(t) = 0 \text{ (eq 6.4, the seismometer at equilibrium)}$$

The ‘poles and zeros’-method (PAZ) refers to a transfer function's points where the function grows to infinity (poles) and where its roots are zero (zeros). Scherbaum (2001) defines the transform function as the ratio of Laplace conversions of the input and output signals, which are similar to the frequency response function $T(j\omega)$, with the exception that $s=j\omega$ is set. The response function of the PAZ-method is

$$H(s) = \frac{B \prod_{n=1}^{N_z} (s - z_n)(s - z_n^*)}{\prod_{k=1}^{N_p} (s - p_k)(s - p_k^*)} \text{ (eq 7),}$$

where N_p is the number of poles, p are the poles, N_z is the number of zeros and z are the zeros, and B is a single value for the sensitivity and damping of the seismogram, or, more broadly, the system (Bormann et al. 2014, Smith and Randall 2016).

In practice, the response removal can be done either by dividing the processed signal with an instrument constant or by using the PAZ-method. Dividing by constant can be used accurately only for frequency range where the instrument response is constant. The PAZ-

approach is better because it simulates the instrument response more accurately. PAZ requires a matrix of response data, the sensitivity and damping of the seismometer, which are provided by manufacturer (Scherbaum 2001, Bormann et al. 2014). Dividing by constant can be used to get approximate values of the actual data. Correlating the two traces, one that has the response removed by dividing with a constant and the other with PAZ, can give good insight into whether the response removal works as intended.

2.4 Fourier Transform and Frequency Domain

Fourier transform is used to transform data from time-domain to frequency domain, and it is used both to simplify certain mathematical processes as well as to study a signal's frequency spectrum (Shakal et al. 2003, Bendat and Piersol 2011). Fourier transformations can be done for continuous and discrete signals, but in the digital world the signals are discrete. As such, instead of using variables t for time and f for frequency, τ is used for time and νN^{-1} in place of frequency, where ν is a frequency related term and N is the number of samples. The discrete Fourier transform for function $f(\tau)$ is (Råde and Westergren 2004):

$$F(\nu) = \frac{1}{N} * \sum_{\tau}^{N-1} f(\tau) e^{-\frac{i2\pi\nu\tau}{N}} \quad (eq\ 8)$$

Equation 8 is for discrete Fourier transform, with $W = e^{-i2\pi}$ expanded. Bendat and Piersol (2011) provide a transform function as well, and with changes of t to τ and f to νN^{-1} and understanding that angular frequency is linked to the true frequency with $\omega = \nu N^{-1}$, the same function is received. For frequency, ν is used instead of f to avoid confusion and to highlight that ν is frequency related.

Differentiation in frequency domain has the advantage of being more accurate than a numerical differentiation, as is shown in Figure 2, and also being mathematically simpler. Fourier transform allows for power spectral density analysis, which is a form of analysis that focuses on how big of a contribution each frequency has to the signal (Bendat and Piersol 2011, Bormann and Wielandt 2013, Bormann et al. 2014). This assumes the signal

to be either infinite length or considered infinite. This is a useful tool for studying proper filtering frequencies for seismic signals. In Figure 3 the left side has the signal's filter as highpass at a frequency of 0.2 Hz, whereas on the right it is at 5.0 Hz. While a seismic event does cause vibration at under 5 Hz frequencies, the level of noise may be stronger, as it is in this case. The filtering frequencies for seismic signals are different at different times. For instance, during daytime anthropogenic noise is stronger than during night-time, so a narrower filter may be required during daytime. Attempting to remove response from a noisy trace will cause the signal to become very noisy, as deconvolution operations tend to enhance the signals at the edges of the frequency spectrum (Scherbaum 2001). Figure 4 shows how an incorrectly filtered data has lost its proper form (green) and looks nothing like the signal it should look like (blue). The high cutoff frequency for local earthquakes tends to be close to the Nyquist frequency of geophones. Nyquist frequency, which is half of the sampling frequency, is used due to the Sampling theorem: a signal measured with a sampling frequency can only be represented properly if the signal has no energy above half of the sampling frequency (Scherbaum 2001, Råde and Westergren 2004). Energy above the Nyquist frequency will cause an aliasing effect. Using either anti-alias measures or a lowpass filter this effect is avoided (Scherbaum 2001). Analog filters in a seismometer can already serve as a lowpass or anti-alias filter for signals above the Nyquist frequency (Asch, 2009).

The derivation of the data in frequency domain uses Equation 9 and integration uses Equation 10 (Scherbaum 2001).

$$\frac{d}{dt} * f(t) \rightarrow j\omega F(j\omega) = F(j\omega) 2\pi i \nu \text{ (eq 9)}$$

$$\int_{-\infty}^t x(\tau) d\tau \rightarrow \frac{1}{j\omega} * X(j\omega) = F(j\omega) / 2\pi i \nu \text{ (eq 10)}$$

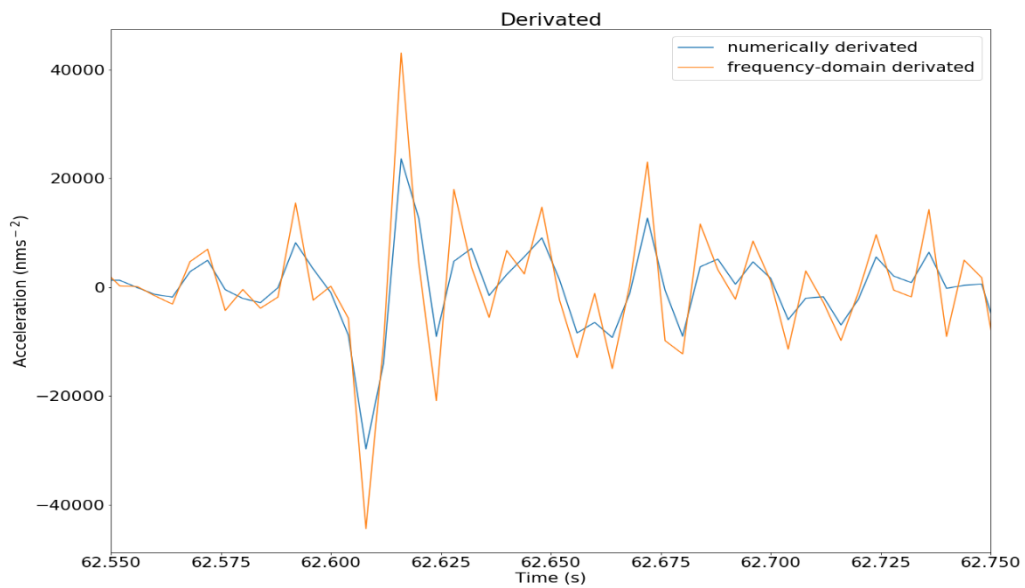


Figure 2. Comparison of ground motion from numerical and frequency domain derivation. Using Fourier transform to receive a frequency domain and inverse Fourier transfer to return to time domain, the signal loses less information than using a straight numerical derivation.

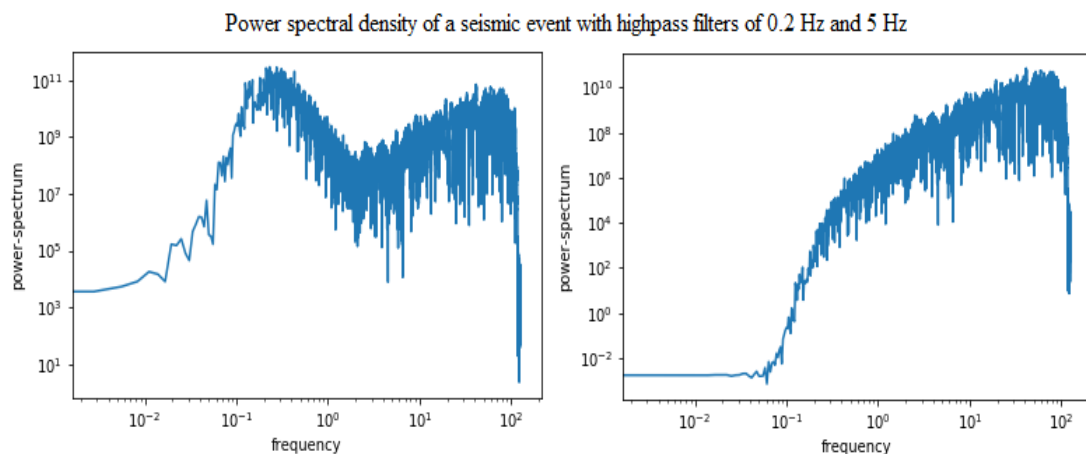


Figure 3. Example of power spectrum from an incorrectly filtered data (left) and a correctly filtered data (right).

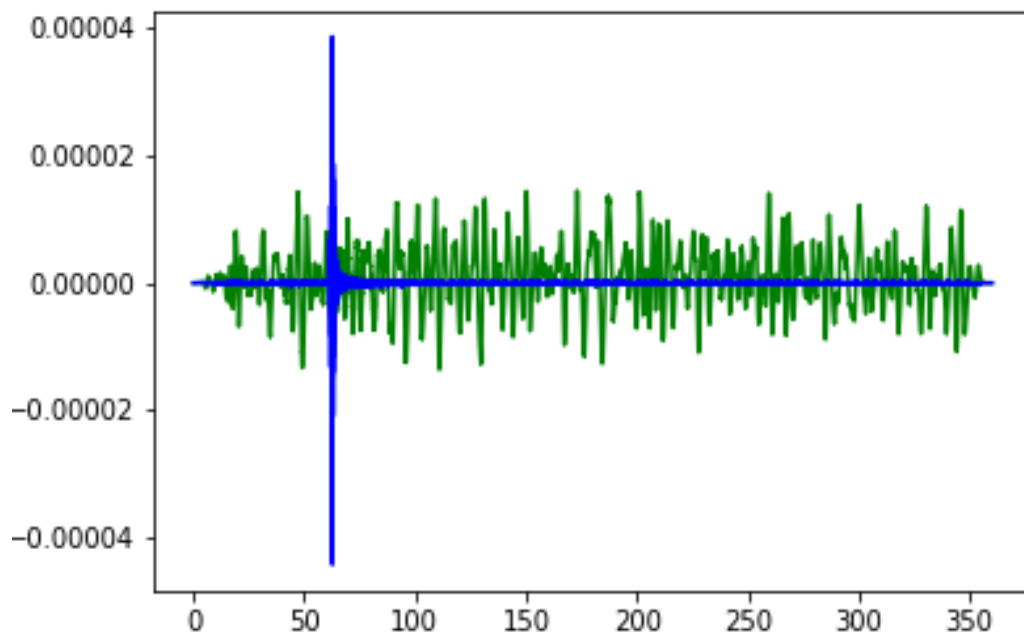


Figure 4. Example of two seismograms: blue is the velocity profile of a correctly filtered data, and green is the velocity profile of an incorrectly filtered data with its values divided by 10^6 to fit both visibly into one picture.

2.5 Ground Motion—From Physical Motion to a Digital Signal

While old seismographs recorded continuous data by drawing the vibrations directly on paper, most modern seismometers measure counts per time and the data they produce is discretized. A modern seismometer has a magnet placed inside of a conductive coil, and as ground movement causes the magnet to move, it induces an electric current into the coil. The induced current causes a voltage, the strength of which depends on the speed of the changing magnetic field. A higher speed induces a larger voltage and in turn a stronger current. The raw data from a modern seismometer is the discretized voltage it measures from the coil (Asch 2009, Wielandt 2012).

The digitizer of a seismometer records the changes in the voltage into a list. Each value in the list is relative to the change in voltage and the elements in the list have an even timestep between them. The frequency of the measurements is given by the manufacturer of a seismometer (e.g. 400 Hz for geophones in this study). Using the frequency and the starting time of the seismogram, one can separately calculate a timestamp for every measurement, although most modern seismometers have in-built clocks and automatically add timestamps to their measurements (Asch 2009).

2.6 Ground Motion–From Digital Signal to Useful Data

Now that the values are on the digital list, everything physical that can affect the value has happened, but the data are not yet in a usable form. The system response is still in the data, and before that can be removed the data requires processing. Three steps are needed: detrending, tapering and filtering.

Plotting seismic raw data tends to result in having apparent first-order trends where the average values of the data either increase or decrease. These first-order trends correspond to high wavelength ground movement and seismometers' internal noise, but in the context of studying an individual seismic event, especially one occurring locally, they are approximated as first order trends. Detrending is the simple operation of removing the trends and fixing the data to a zero level, or, in other words, to average the movement into zero (Bendat and Piersol 2011, Wielandt 2012).

Tapering affects the beginning and end of the data in question. For this study, the tapering was 5% of the trace length from the beginning and ending, so for a trace of 40 s, 2 seconds were tapered from both ends. Tapering is done for the use of frequency domain and is an important step for Fourier transformations to work later (Bendat and Piersol 2011, Bormann et al. 2014).

Next is the instrument response removal. Python's ObsPy library is made specifically for processing seismic data and provides ready-made functions for the PAZ-method which

requires an input of the instrument response parameters (Beyreuther et al. 2010). Once the instrument response is removed, the actual ground motion is left. Depending on the seismometer the data can be in different forms as either displacement, velocity, or acceleration data. As the components are derivatives and integrals of each other, one can calculate the other two either numerically or more precisely using the frequency domain through the Fourier transform (Bormann and Wielandt 2013). Once the signal has been differentiated, an inverse Fourier transform returns it to time-domain.

3. DATASET

The dataset that is used to create the GMPEs and what is used in studying the effect of azimuth is from Otaniemi EGSs borehole OTN-III stimulation of 2018. Table 1 is a summary of the dataset. Figure 5 shows a map of all the stations and temporary geophone arrays that were used during the stimulation (Hillers et al. 2020). Figure 6 shows the magnitude distribution of the events. Figure 7 shows the epicentral locations of the events relative to each other, plotted with their depth and magnitude and Figure 8 the magnitudes plotted with the depth of the hypocentres. The depths of the events are in three separate clusters: the upper, the central and the bottom cluster (Leonhardt et al. 2021).

Table 1. The dataset of the 2018 stimulation earthquakes.

2018 stimulation	
DOI: 10.5880/GIPP.201802.1	
number of events	207 of which 204 are used
magnitude range	0.0–1.8
depth range	4800–6300 m (one outlier at 10 300 m)
distance to station range	(manual cutoff at 20 km)
Total number of records	20 768
ISUH stations' records	17 829
ST1 stations' records	2 427

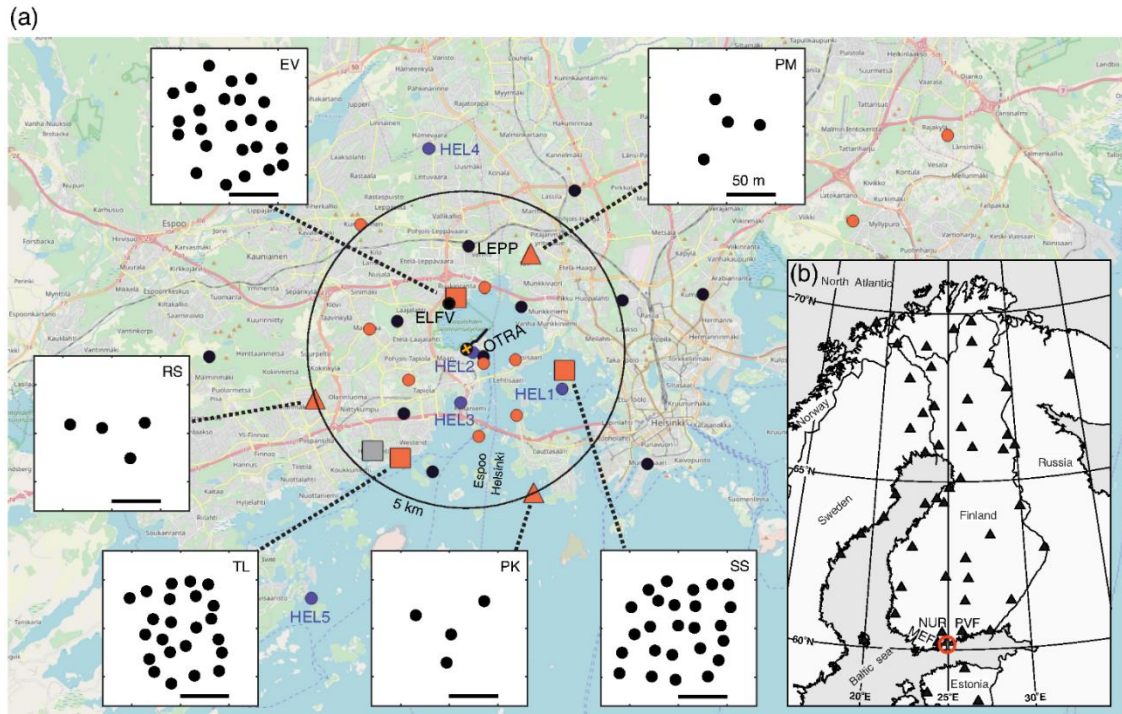


Figure 5. a) Map of the surface and satellite stations in use during the 2018 stimulation. HEL-broadband network is shown as blue circles, the orange symbols are for all the cube-stations and arrays, and black circles are for borehole stations. The orange 'x' marks the injection site at Otaniemi, with the black trail marking the direction of the borehole; b) Map of Finland showing the national seismic station network, with the closest stations MEF, NUR and PVF also receiving recordings from the largest events (Hillers et al. 2020)

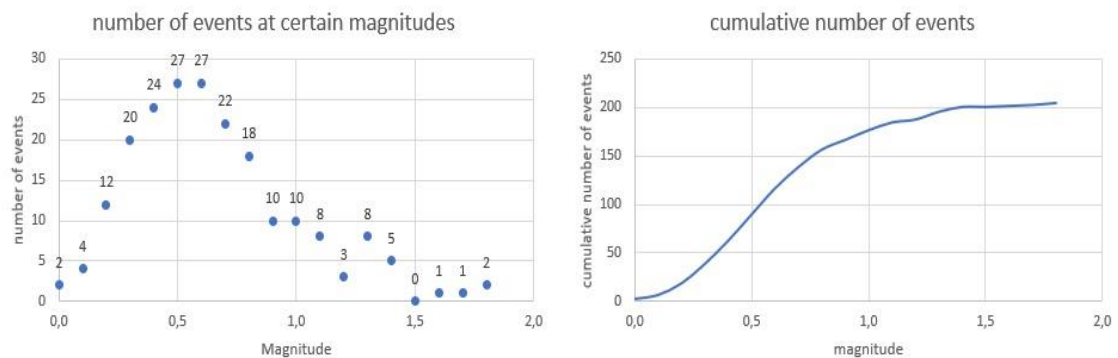


Figure 6. Magnitude distribution (left) and cumulative magnitude distribution (right) of the events in the database used to compile a peak ground motion database.

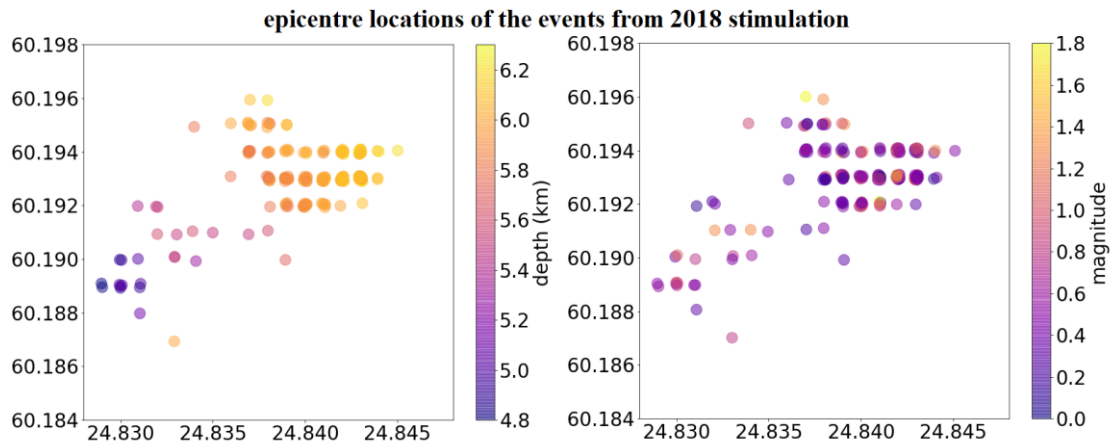


Figure 8. Epicentre distribution of the events relative to each other with depth (left) and magnitude (right) with different colours. Since the events have a 3-decimal location in WGS84 coordinate system, plotting them results in a lot of the data becoming invisible. Each point was given a small random number to their N and E coordinates to aid in visualizing the data.

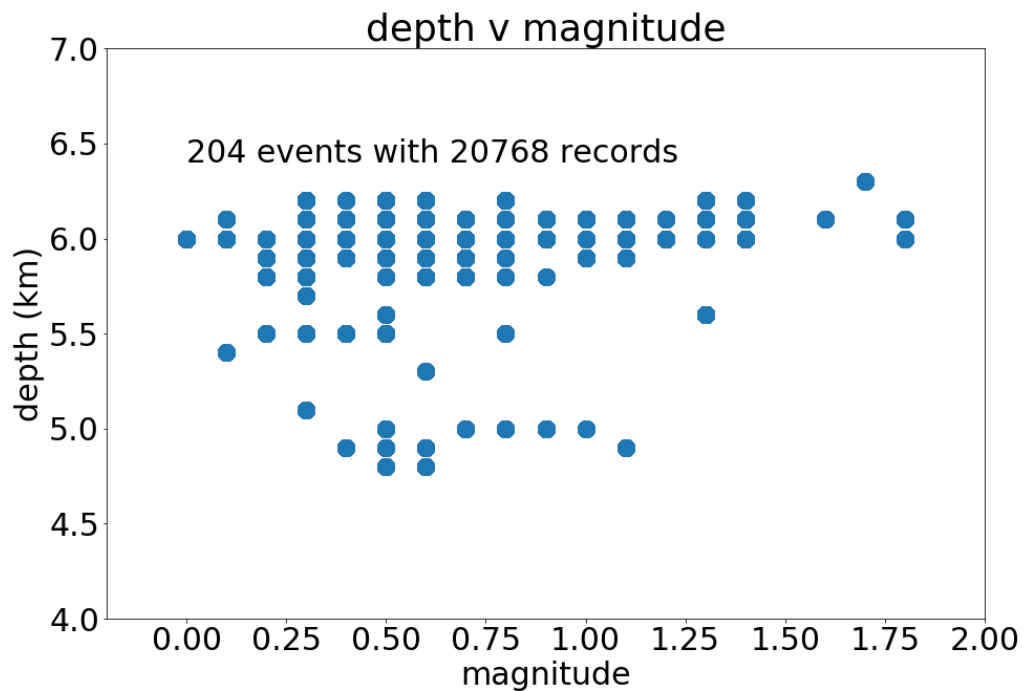


Figure 7. Depth of the hypocentre relative to the magnitude of the event. The three clusters as mentioned by Leonhardt et al. (2021) can be seen.

During the 2018 stimulation, due to the number of seismic events from the borehole the analysts were instructed to prioritize big events, and after the stimulation ended and the seismicity dropped, more events were analysed, as well as the previously analysed events were analysed in more detail, with more stations in use. As a result, most of the events of the dataset have magnitudes close to 0.5. Events around and below magnitude 0.0 did occur, but time constraints prevented their full analysis. Figure 9 shows a few examples of fault-mechanism solutions (Hillers et al. 2020). Considering most solutions were reverse-fault mechanisms and the only strike-slip solution is considered less reliable (Hillers et al. 2020), this study will not consider the effect of the mechanism.

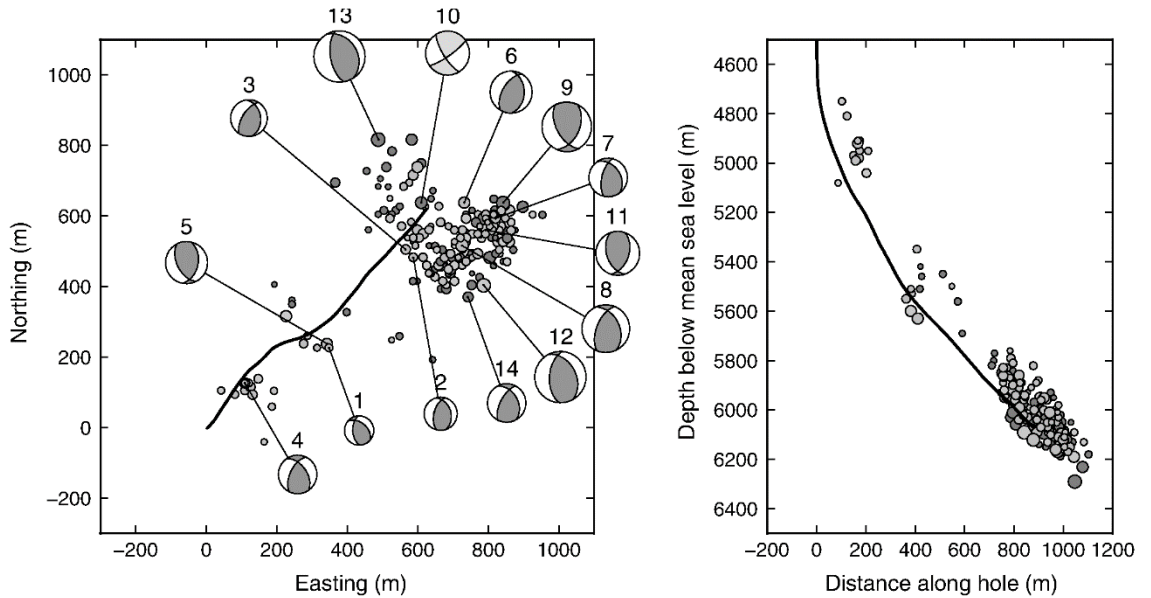


Figure 9. A few examples of fault-mechanism solutions (left) and the depth distribution of events with the borehole as black (right; modified, original: Hillers et al. 2020).

4. METHODS—FROM RAW DATA TO A MODEL

From raw data to a functional GMPE three main steps were required: gathering of the peak ground motion data, the creation of the initial models, and solving for the constants.

The first step of gathering the peak ground motion data was done in a single Python program script, with the use of ObsPy-library. While the ObsPy-library has ready-made codes for the processing of seismic data, a function was made that took the processed data and differentiated or integrated it in frequency domain (Appendix 1). The peak ground motion values were maximum absolute values from the processed data. For horizontal values, the vector length of the combined x- and y-axis components was calculated. The finished peak ground motion dataset was saved as a CSV-file for convenient use later (an example in Appendix 2, full dataset available from author). A flowchart of this process is shown in Figure 10.

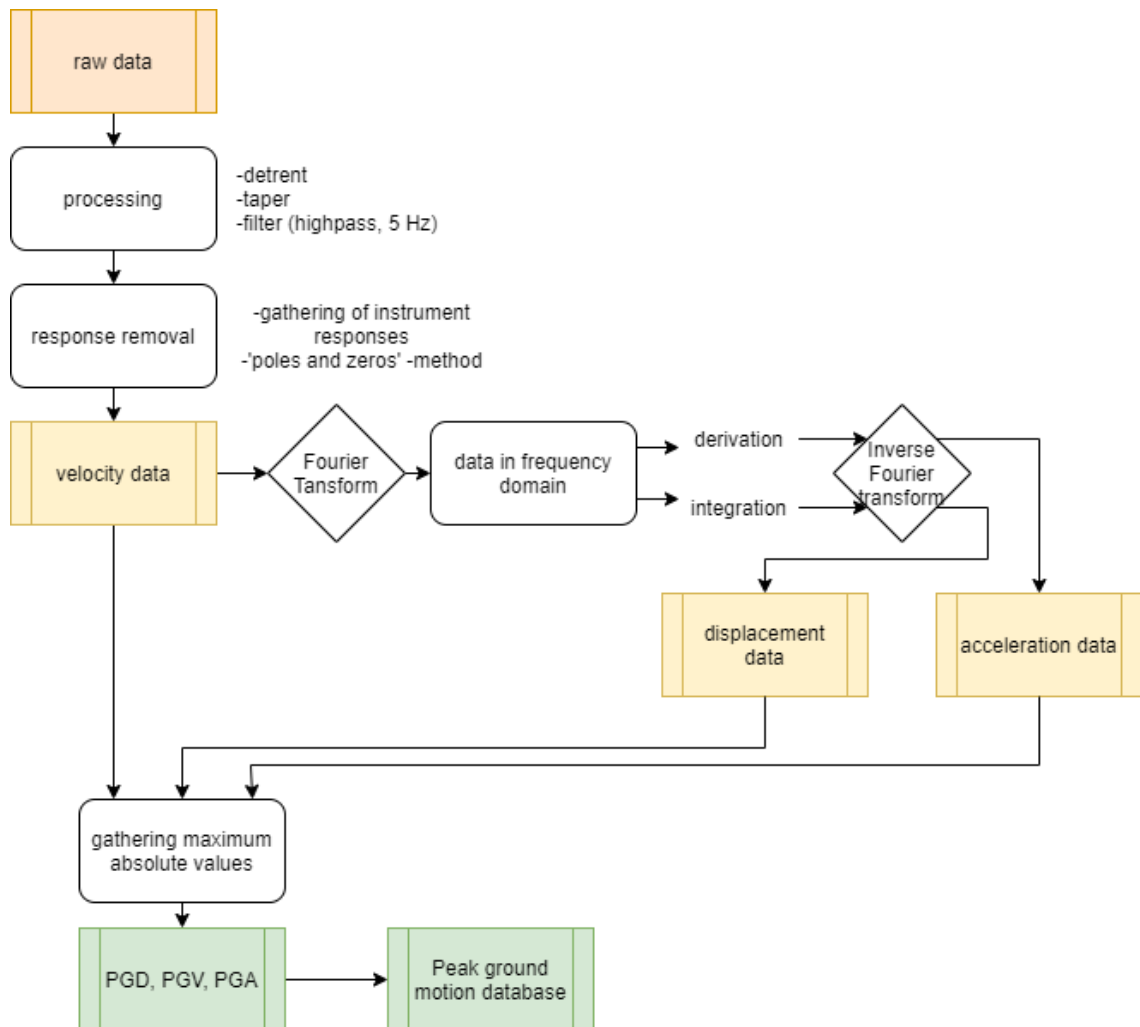


Figure 10. Flowchart of the process from raw data to a peak ground motion database.

After the peak ground motion dataset was compiled, it was used to create the GMPE models. The models were created by fitting modified Campbell's (2003) model (Equation

3) into the peak ground velocity data one magnitude's events at a time, with distance from the hypocentre as the only variable. Before the fitting, the full dataset was filtered by magnitude and maximum hypocentral distance. The hypocentral distance was set at $r < 20$ km for the models, while each fit had a different magnitude. By constricting the magnitude and distance the constants received at different magnitudes varied considerably less than if there were no distance filters. The fitting of the filtered peak ground motion data (vertical and horizontal PGV and PGA) was done using Python's SciPy-module's `curve_fit()`-function, which is a non-linear least-squares fitting function that uses a Levenberg-Marquardt algorithm (Moré et al. 1980, SciPy v1.6.2 Reference Guide 2021). Because the GMPE's are regressions fits (Campbell 2003), they ultimately require testing to get the best fits. All datasets were fitted at least four times with tightening bounds for the constants. The fitting process of vertical PGV GMPE's constants is used as an example of the fitting processes next.

Each fitting started with setting initial bounds to prevent unreasonable values (Bounds 1), and after that they were constricted to better the fit, but not to the point where the constants' values would have been the same as their bounds.

Once the fitting provided constants c_{1-3} (see Equation 3) for every other tenth of magnitude (from 0.0 to 1.8 with 0.2 steps), the constants were plotted with the magnitude (Figure 11a). The constants seemed to follow apparent first order trends, and while in some cases there were indication that the trends may be logarithmic or sinusoidal, at small distances ($r < 20$ km) and magnitudes ($M_L < 2.0$) they were assumed to obey first-order trends to simplify their solving. Shallow trends indicated reasonable average values for the constants, and once the standard deviation of the constants was less than 0.1, they provided decent values for the model.

Constant c_3 , which multiplies the distance term r , had the lowest standard deviation upon initial fitting at different magnitudes (Figure 11b), and thus its value was set as the average of what `curve_fit()` calculated with different magnitudes. After fixing its value, the curve fitting was repeated with just two constants c_1 and c_2 and tightened their bounds

based on their values from initial fitting (Bounds 2). Because the values received for every other one tenth of magnitude had shallow trends and low standard deviation (Figure 11c), new bounds were set for all three constants (Bounds 3) and fitting them at every tenth of a magnitude was attempted (Figure 12a).

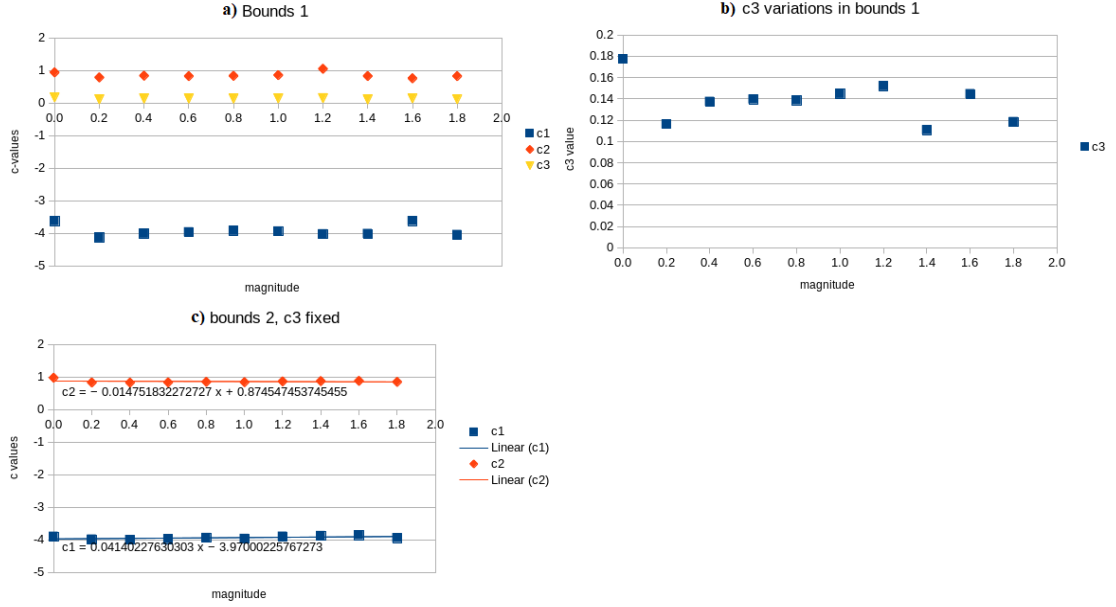


Figure 11. a) Received constants with bounds 1; b) the c_3 variations with bounds 1; c) after fixing c_3 to its average value received with bounds 1, the bounds 2 were used to plot the constants c_1 and c_2 . Because the constants are seemingly unaffected by magnitude changes, every one tenth of magnitude are used (in Figure 12), instead of every other one tenth between 0.0 and 1.8.

After receiving a low standard deviation for c_3 at different magnitudes (Figure 12b), its value was fixed at its average. The final bounds for c_1 and c_2 were then set (Bounds 4) and the values they received from `curve_fit()` were plotted. The average values of the constants and their standard deviations were compared to the range of their bounds (Figures 12c, d, and e). Because the bounds were not touched in the fitting, and their standard deviations were low, their averages were accepted as the final values. If the constants' values touched the bounds, the bounds were widened. Table 2 shows the final bounds' values. Figure 13 illustrates the process from the peak ground motion values into a GMPE.

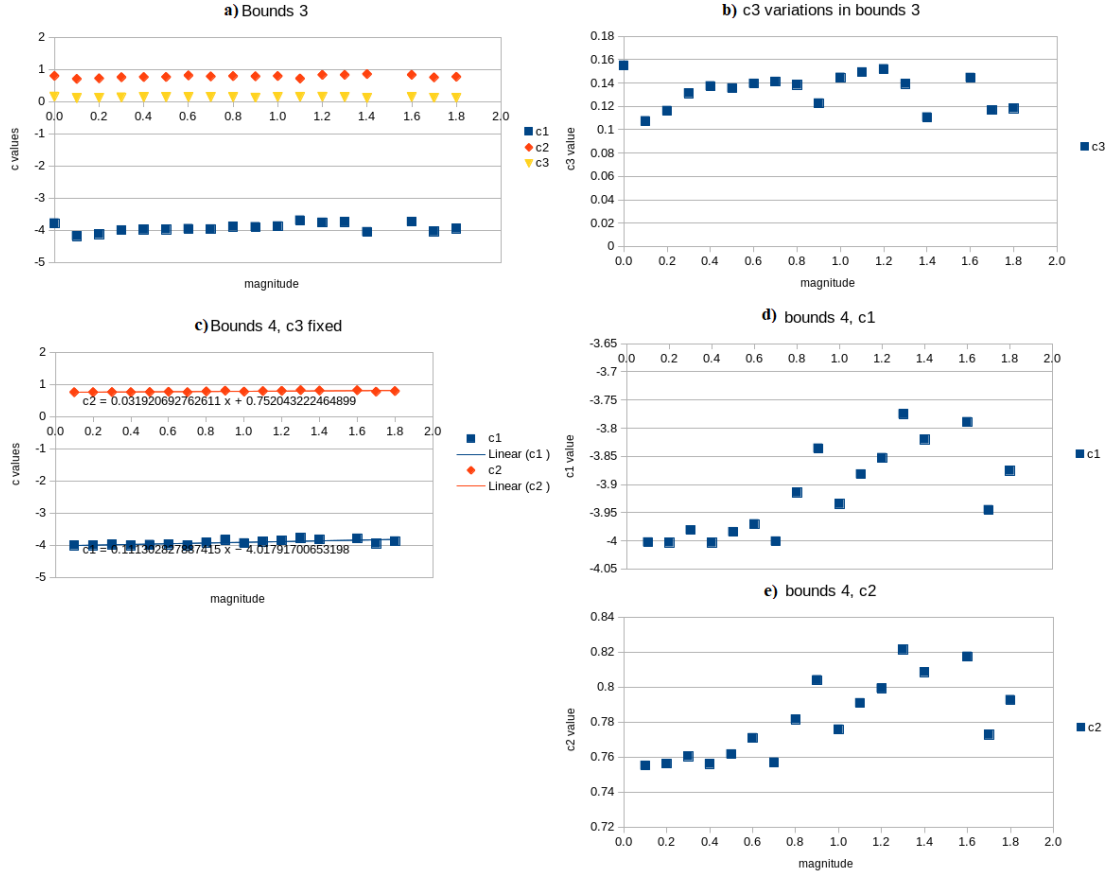


Figure 12. a) All three constants plotted with bounds 3 for every one tenth of magnitude from 0.0 to 1.8; b) c_3 variation with bounds 3. While there is some variation, the standard deviation was so low that an average value was justifiable to use; c) c_1 and c_2 in bounds 4 for every one tenth of magnitude; d) c_1 variation within bounds 4; e) c_2 variation within bounds 4. While c_1 and c_2 seem to show some evolution with magnitude, their deviations were so low that their average values were justifiable to use, although this may diminish the accuracy of the model at $M_L > 1.0$. Table 4, with the summary of the new model ON21 in section 5. RESULTS, has the standard deviations marked next to the constants' values.

Table 2. Bounds used for constants' variation in curve_fit at different stages for the new model ON21.

PGV	c1		c2		c3	
	vertical	horizontal	vertical	horizontal	vertical	horizontal
Bounds 1	-6, 6	-6, 6	0, 5	0, 5	0, 5	0, 5

Bounds 2	-4.2, -3.0	-4.2, -1.0	0.75, 1.2	0.75, 1.0	fixed	fixed
Bounds 3	-4.2, -3.5	-4.2, -3.5	0.70, 0.90	0.70, 0.85	0.105, 0.155	0.105, 0.155
Bounds 4	-4.2, -3.5	-4.2, -3.5	0.70, 0.90	0.70, 0.85	fixed	fixed
PGA	c1		c2		c3	
	vertical	horizontal	vertical	horizontal	vertical	horizontal
Bounds 1	-5, 5	-5, 5	0, 5	0, 5	0, 5	0, 5
Bounds 2	-2, 0	-2.5, -0.5	0.7, 1	0.8, 2.5	fixed	fixed
Bounds 3	-2, 0	-4, -1.5	0.7, 1	0.5, 1.2	0.13, 0.195	0.12, 0.2
Bounds 4	-2, 0	-4, -1.5	0.7, 1	0.5, 1.2	fixed	fixed

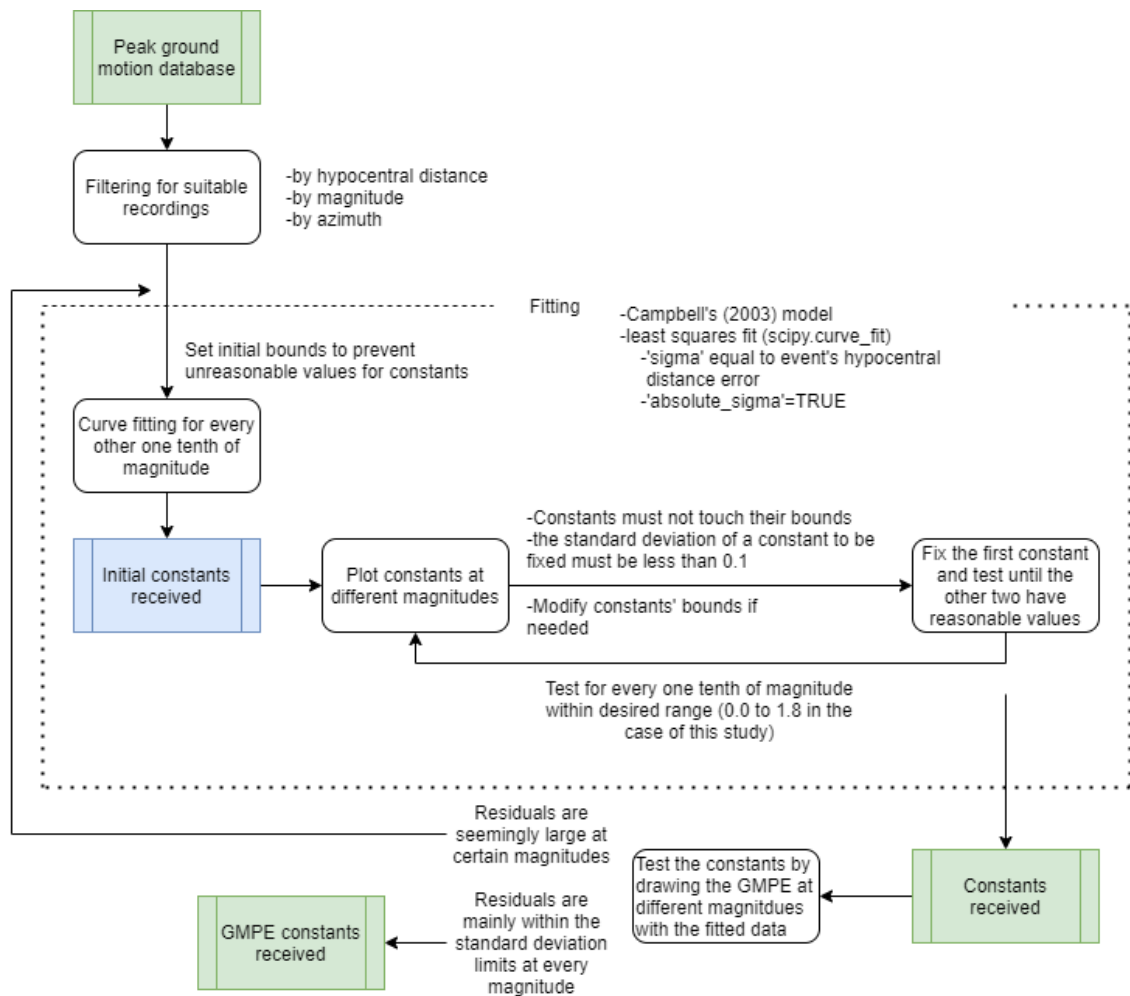


Figure 13. A flowchart of the process from PGV values to the constants of the GMPE model.

5. RESULTS

The peak ground motion database from the 2018 stimulation data was compiled to a CSV-file (an example in Appendix 2, full dataset available from author). Table 3 summarizes what information each column of the CSV-file has. Figure 14 shows how many records are from which arrays. Due to the limitation of 20 km, the records from the accelerometers of the national Finnish seismic network were not used. Most records came from arrays with the most stations; Elfvik-array had 24 geophones whereas DT-network had only 4 stations, meaning Elfvik could record the same event more times. Figure 15 shows the maximum, mean, standard deviation and median values' evolution by magnitude for the peak ground motions.

Table 3. Summary of the peak ground motion database.

st1-18_database_PGM.csv			
contains a total of 20768 records of 204 events			
column name	explanation	data type	example
id	event id as in database	integer	195078
station	the station of the record	string	TL15
network	the network of the station	string	Toppelund_cubes
M	magnitude	float	0.6
M_error	magnitude error	float	0.2
depth(m)	hypocentre depth of the event	float	5800.0
depth_error(m)	error of the depth of the event	float	0.0
fixed_depth	whether the depth was fixed, forcing the depth_error to 0	boolean	True
distance(m)	hypocentral distance	float	7488.88751882
event_lat	latitude of the event in WGS84 coordinate system	float	60.193
lat_error(m)	latitude error of the event	float	10.0
event_lon	longitude of the event in WGS84 coordinate system	float	24.838

lon_error	longitude error of the event	float	10.0
PGD(mm)	vertical peak ground displacement	float	0.000125907905229
PGV(mm/s)	vertical peak ground velocity	float	0.0438337525065
PGA(mm/s ²)	vertical peak ground acceleration	float	31.2020316687
filtering	what filter was used	list	“['highpass', 5.0]”
azimuth	the (front)azimuth of the recording to its event (in degrees clockwise, 0 ° is north)	integer	235
PGD_hor(mm)	horizontal peak ground displacement	float	0.000178060667185
PGV_hor(mm/s)	horizontal peak ground velocity	float	0.0678611807402
PGA_hor(mm/s ²)	horizontal peak ground acceleration	float	44.1263363594
statlon	the longitude of the station	float	24.78816
statlat	the latitude of the station	float	60.15846

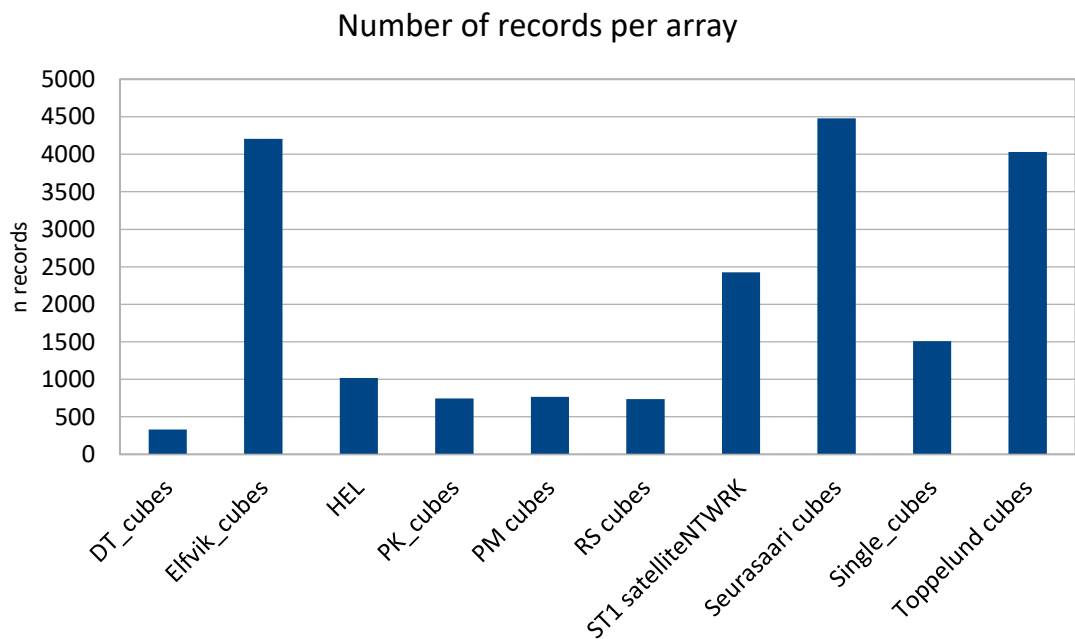


Figure 14. Distribution of records from events of 2018 stimulation from different arrays and networks.

Appendix 3 has a table of the maximum, mean, standard deviation and median values of vertical PGD (in mm), PGV (in mms^{-1}) and PGA (in mms^{-2}) that were calculated, and Figure 15's data is from the table of Appendix 3. The biggest values for all ground motion came from large events ($M_L = 1.8$) at the closest stations, and the minimum values came from the stations farthest from the hypocentre. Seurasaari array and MURA of ST1 satellite network had the highest values in all peak ground motions.

There is a clear strengthening of ground motion as the magnitude increases with some exceptions. At $M_L = 0.0$ the PGV seems to be higher on average than at $M_L = 0.1$ and at 1.7 the average PGV seems to fall. Both may be results of a data-bias: because less $M_L = 0.0$ events were analysed than for instance $M_L = 0.1$, the standard deviation of $M_L = 0.0$ is higher than that of $M_L = 0.1$. The small standard deviation at $M_L = 1.7$ is due to there having been but one $M_L = 1.7$ event in the database, but two $M_L = 1.8$ events.

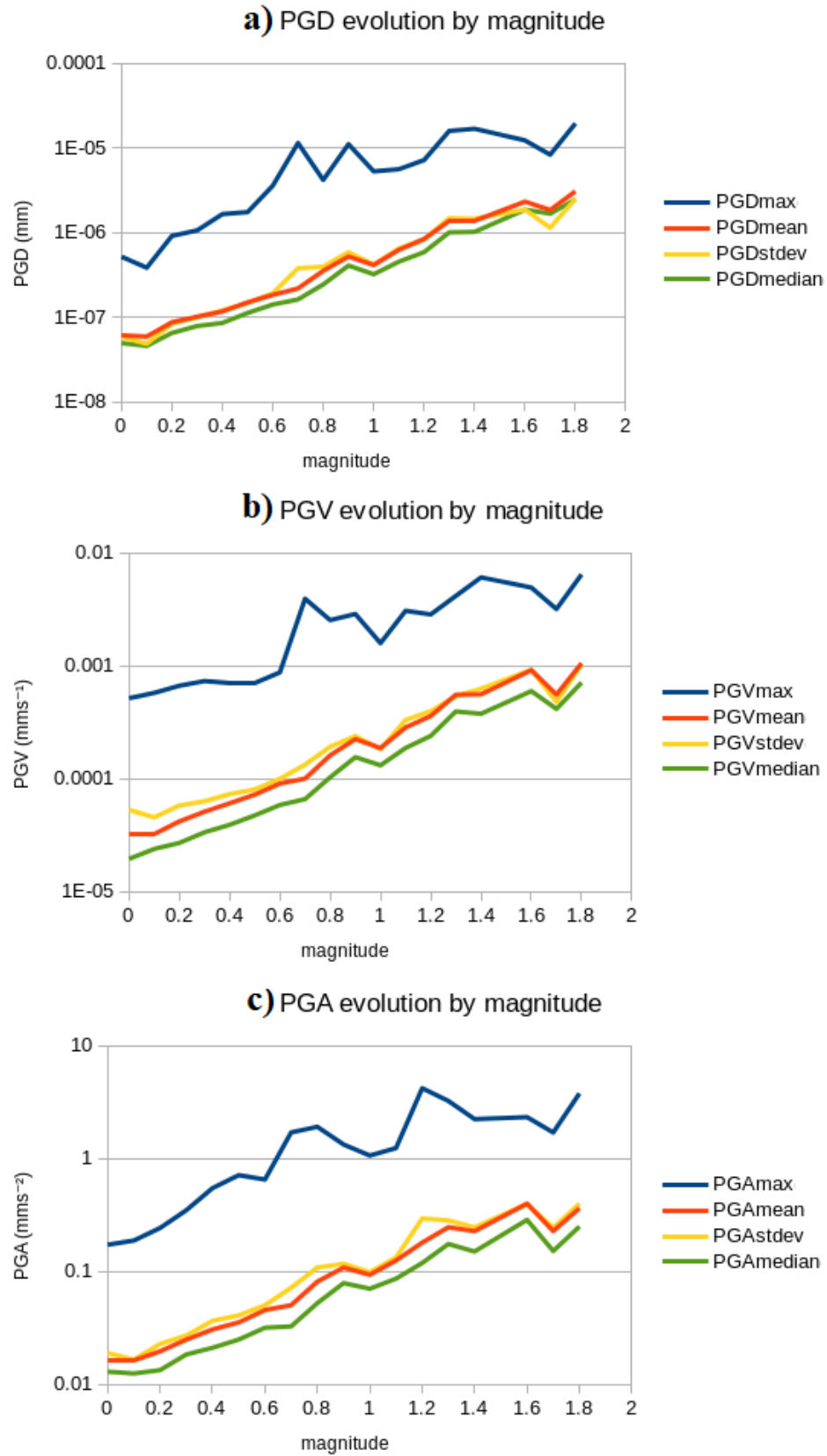


Figure 15. Evolution of maximum, mean, standard deviation and median of peak ground motion values for vertical a) PGD, b) PGV and c) PGA on a common logarithmic scale.

Table 4 shows the final model for peak ground velocity, named as ON21 (for OtaNiemi 2021) for both vertical and horizontal ground velocity and acceleration. Figure 16 shows some examples of the model at different magnitudes for vertical PGV and Figure 17 for horizontal PGV, Figure 18 for vertical PGA and Figure 19 for horizontal PGA. The red line is the model, and the blue lines are its 1σ ranges. The model ON21 1σ -uncertainty for vertical PGV is 0.60, which is lower than the 0.81 of Douglas et al.'s (2013) model and 0.62 of Fin17 by ISUH that were used initially (according to Ader et al. 2020). Model ON21's 1σ -value for horizontal PGV is 0.68, indicating more deviation in horizontal than vertical ground motion. Similarly, for vertical PGA the 1σ -value is 0.61 and for horizontal PGA it is 0.64. Because the PGA values were received by taking the maximum absolute value from derivated velocity data, instead of having the measurements directly as acceleration, there is more variation with the data than in the velocity data. Figure 20 shows measurements at two different magnitudes, with different arrays coloured. The ST1 satellite network's stations (pink dots) seem to on average have lower values for peak ground velocity than the ISUH stations.

Table 4. A summary of the solved model ON21 for PGV and PGA in vertical and horizontal ground motions.

GMPE model ON21		
Magnitude range		0.0–1.8
Distance range		up to 20 km
for peak ground velocity (PGV)		
c ₁	vertical	-3.916 ± 0.080
	horizontal	-3.925 ± 0.087
c ₂	vertical	0.781 ± 0.023
	horizontal	0.786 ± 0.013
c ₃	vertical	0.133 ± 0.015
	horizontal	0.137 ± 0.014
$\sigma_{\log 10(\text{PGV})}$	vertical	0.598
	horizontal	0.676
for peak ground acceleration (PGA)		
c ₁	vertical	-1.099 ± 0.067

	horizontal	-1.235 \pm 0.078
c ₂	vertical	0.836 \pm 0.011
	horizontal	0.991 \pm 0.020
c ₃	vertical	0.153 \pm 0.017
	horizontal	0.153 \pm 0.017
$\sigma_{\log 10(PGA)}$	vertical	0.611
	horizontal	0.642
$\log_{10}(PGV_v) = -3.916 + 0.781 * M_l - 0.133 * r \pm 0.598$ (eq 11)		
$\log_{10}(PGV_h) = -3.925 + 0.786 * M_l - 0.137 * r \pm 0.676$ (eq 12)		
$\log_{10}(PGA_v) = -1.099 + 0.836 * M_l - 0.153 * r \pm 0.611$ (eq 13)		
$\log_{10}(PGA_h) = -1.235 + 0.991 * M_l - 0.153 * r \pm 0.642$ (eq 14)		

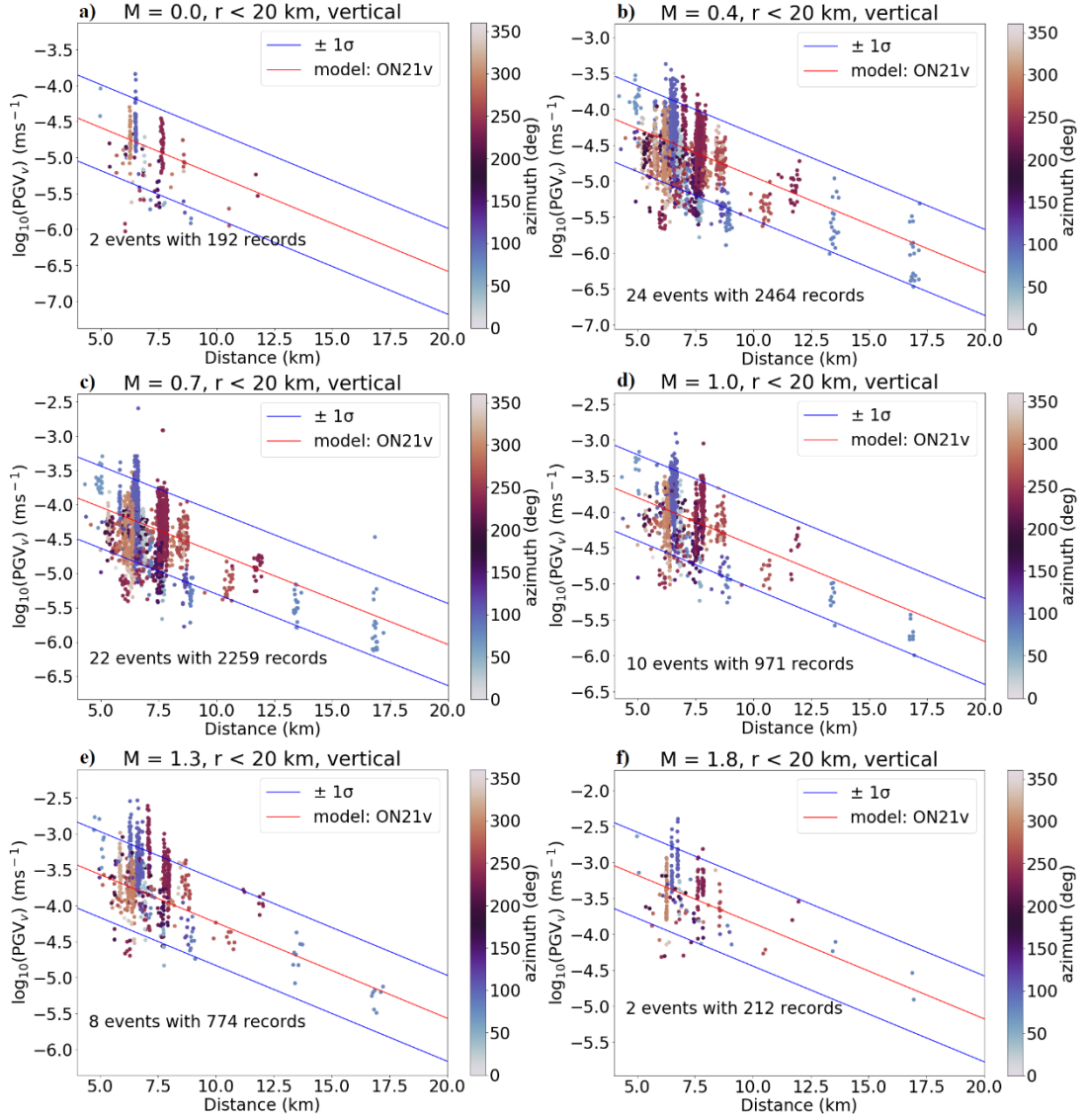


Figure 16. Examples of the model ON21 GMPE for vertical PGV at magnitudes a) 0.0, b) 0.4, c) 0.7, d) 1.0, e) 1.3 and e) 1.8. The red lines are the model, blue lines are the model at $\pm 1\sigma$ ranges and the colours of the data points refer to their azimuth from north direction in a clockwise rotation.

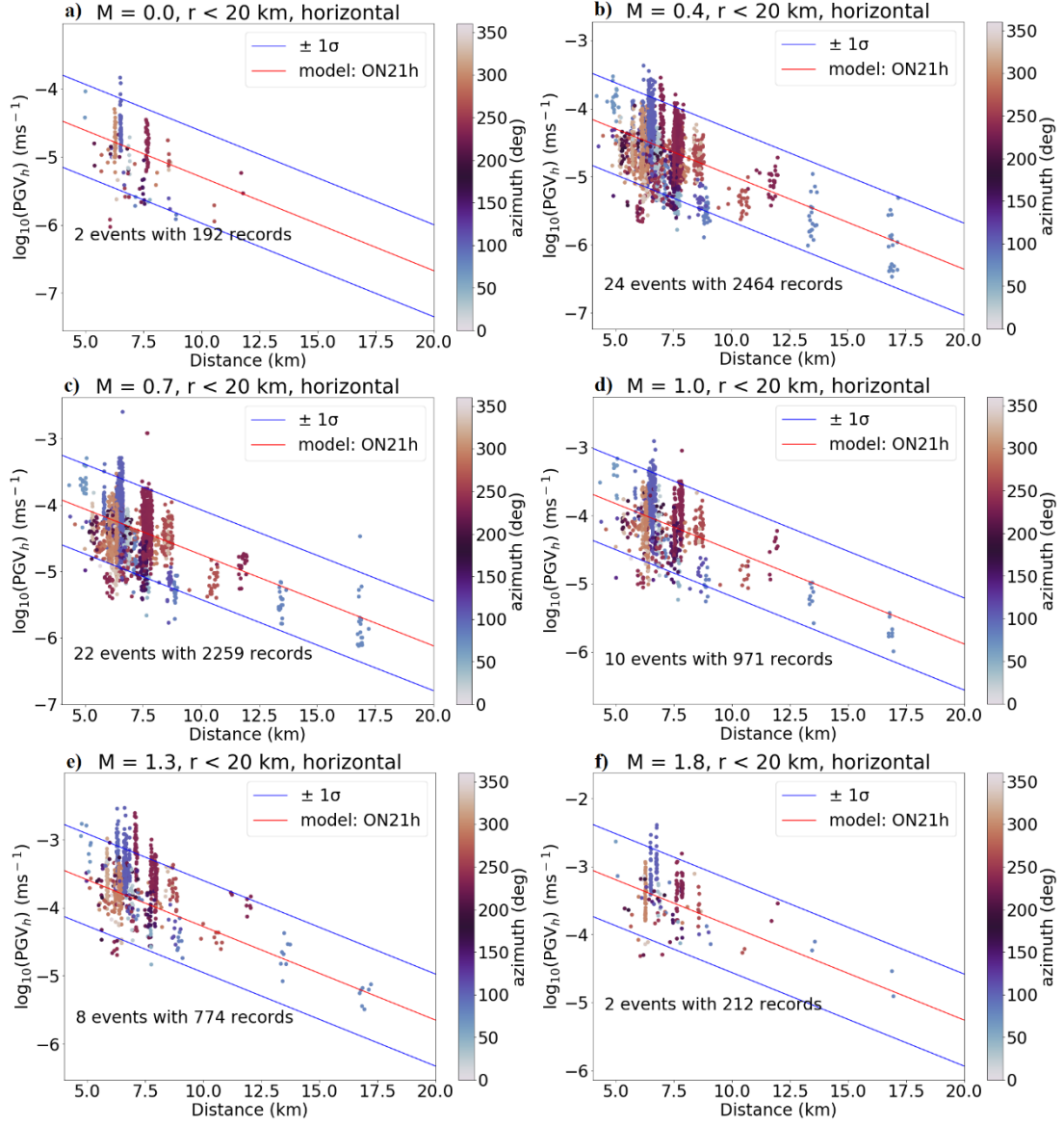


Figure 17. Examples of the model ON21 GMPE for horizontal PGV at magnitudes a) 0.0, b) 0.4, c) 0.7, d) 1.0, e) 1.3 and e) 1.8. The red lines are the model, blue lines are the model at $\pm 1\sigma$ ranges and the colours of the data points refer to their azimuth from north direction in a clockwise rotation.

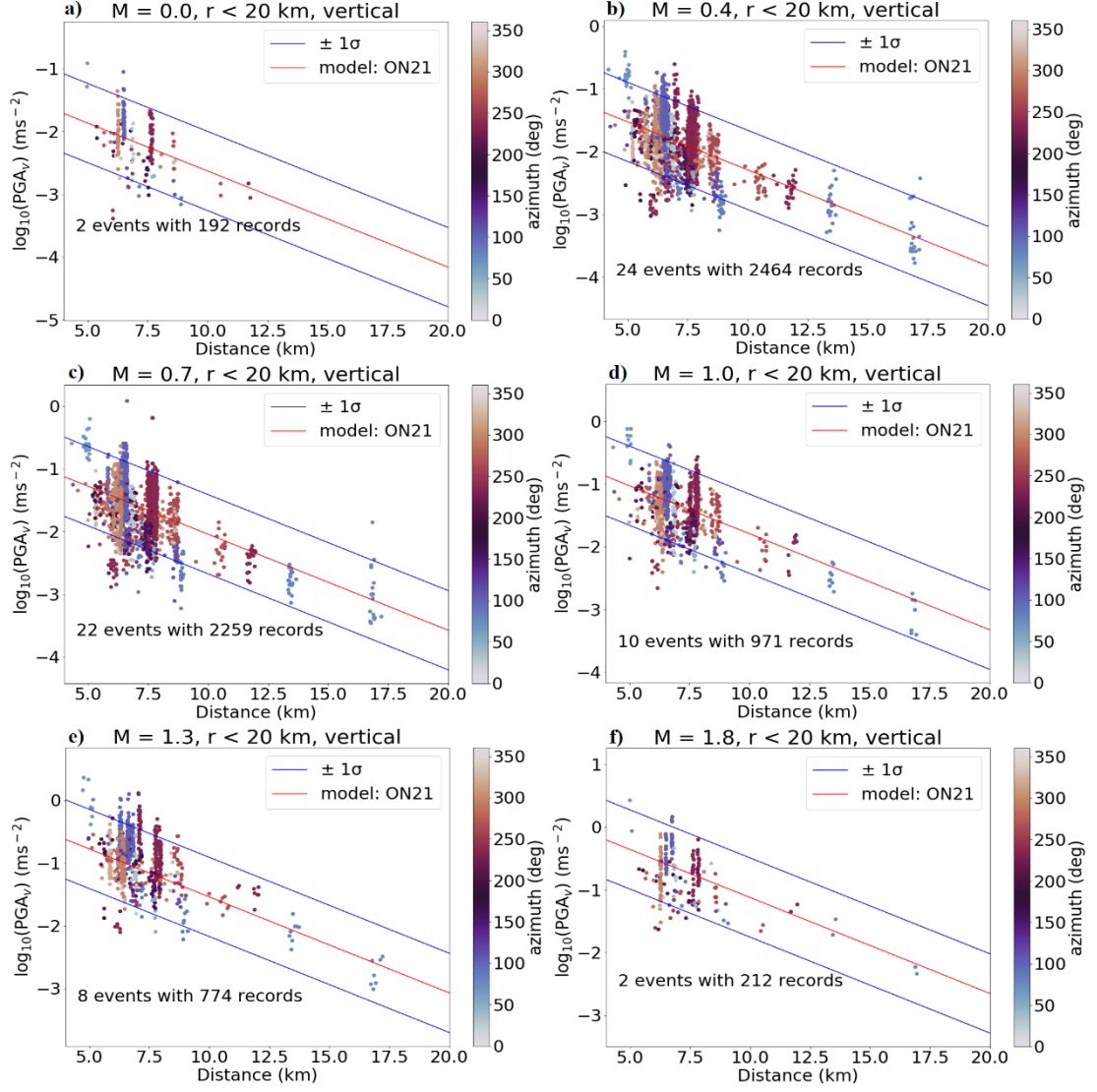


Figure 18. Examples of the model ON21 GMPE for vertical PGA at magnitudes a) 0.0, b) 0.4, c) 0.7, d) 1.0, e) 1.3 and e) 1.8. The red lines are the model, blue lines are the model at $\pm 1\sigma$ ranges and the colours of the data points refer to their azimuth from north direction in a clockwise rotation

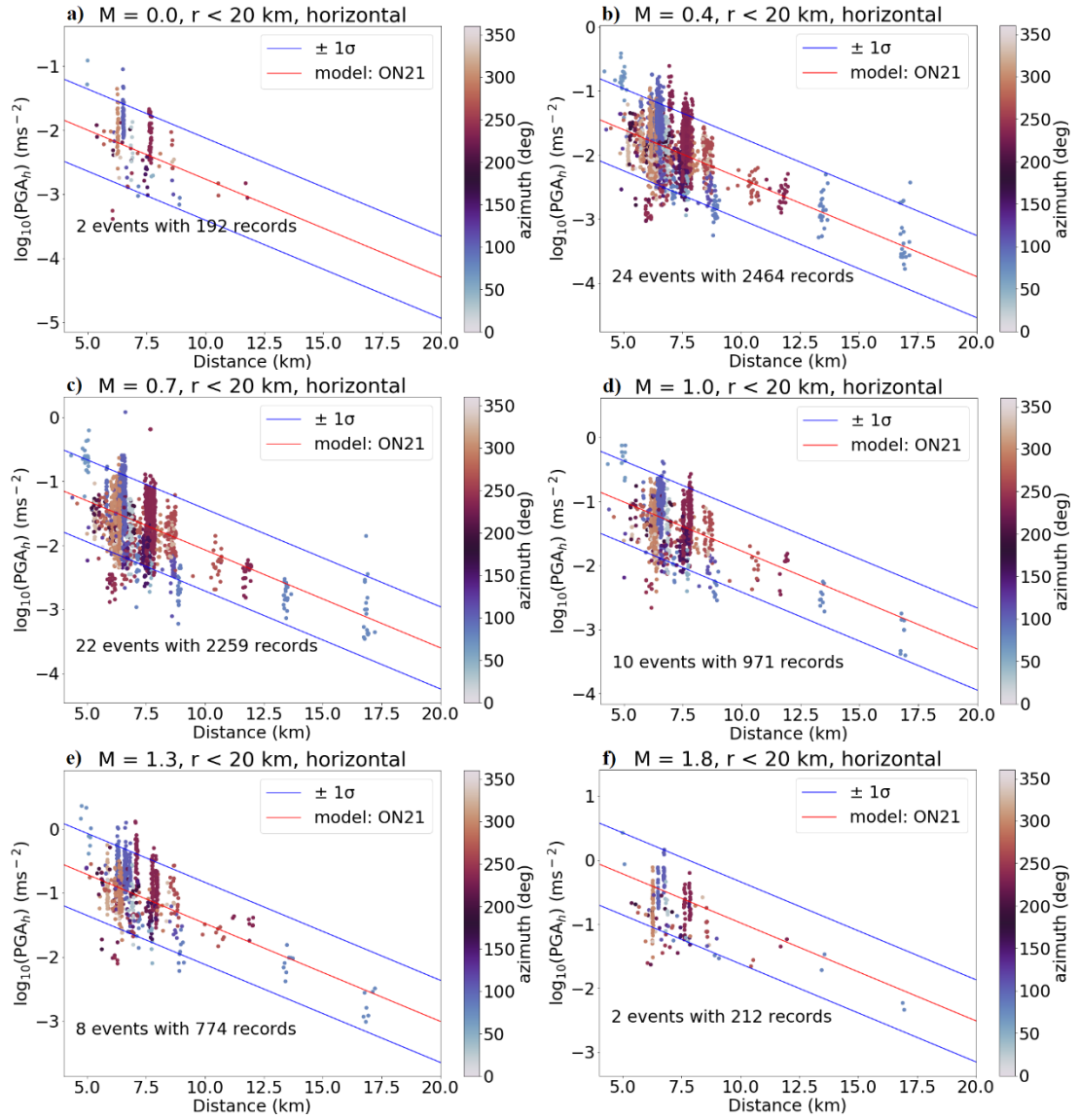


Figure 19. Examples of the model ON21 GMPE for horizontal PGA at magnitudes a) 0.0, b) 0.4, c) 0.7, d) 1.0, e) 1.3 and e) 1.8. The red lines are the model, blue lines are the model at $\pm 1\sigma$ ranges and the colours of the data points refer to their azimuth from north direction in a clockwise rotation

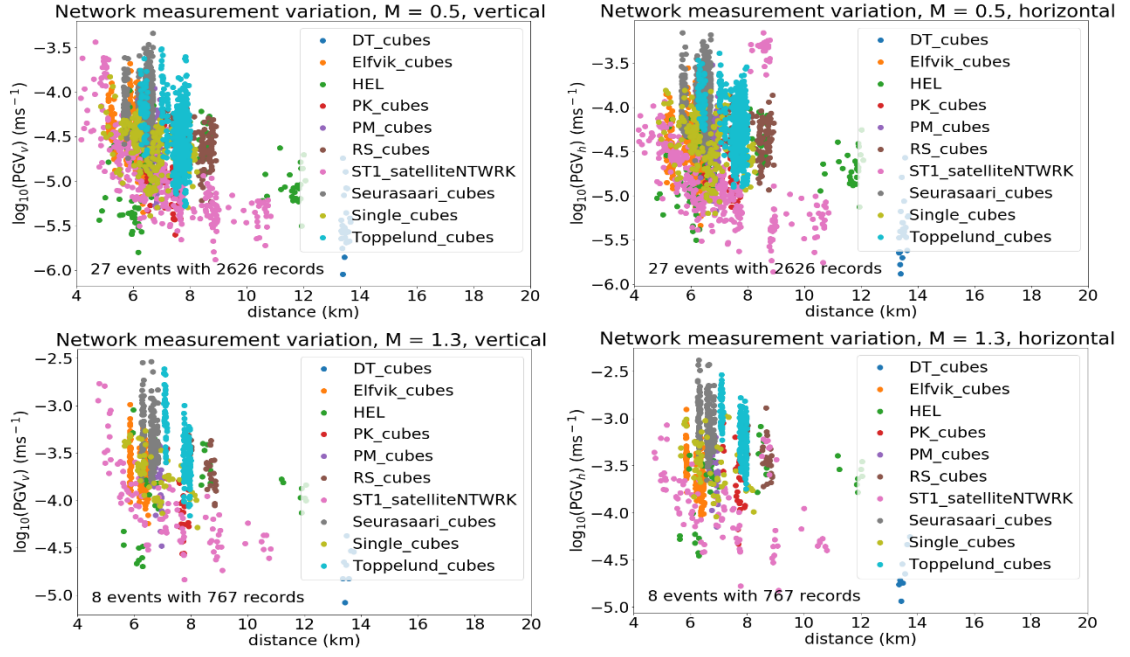


Figure 20. The PGV values at $M_l = 0.5$ (up) and $M_l = 1.3$ (lower) for vertical (left) and horizontal (right) ground motion with different networks at different colours.

Figure 21 shows the azimuth variation with all magnitudes plotted from stations at 2 km intervals for vertical PGV, and Figure 22 for horizontal PGV, figure 23 for vertical PGA and Figure 24 for horizontal PGA. Figure 25 shows four different events with their average recorded maximum PGV on a common logarithm from multiple-station arrays and the maximum PGV recorded at single-station arrays.

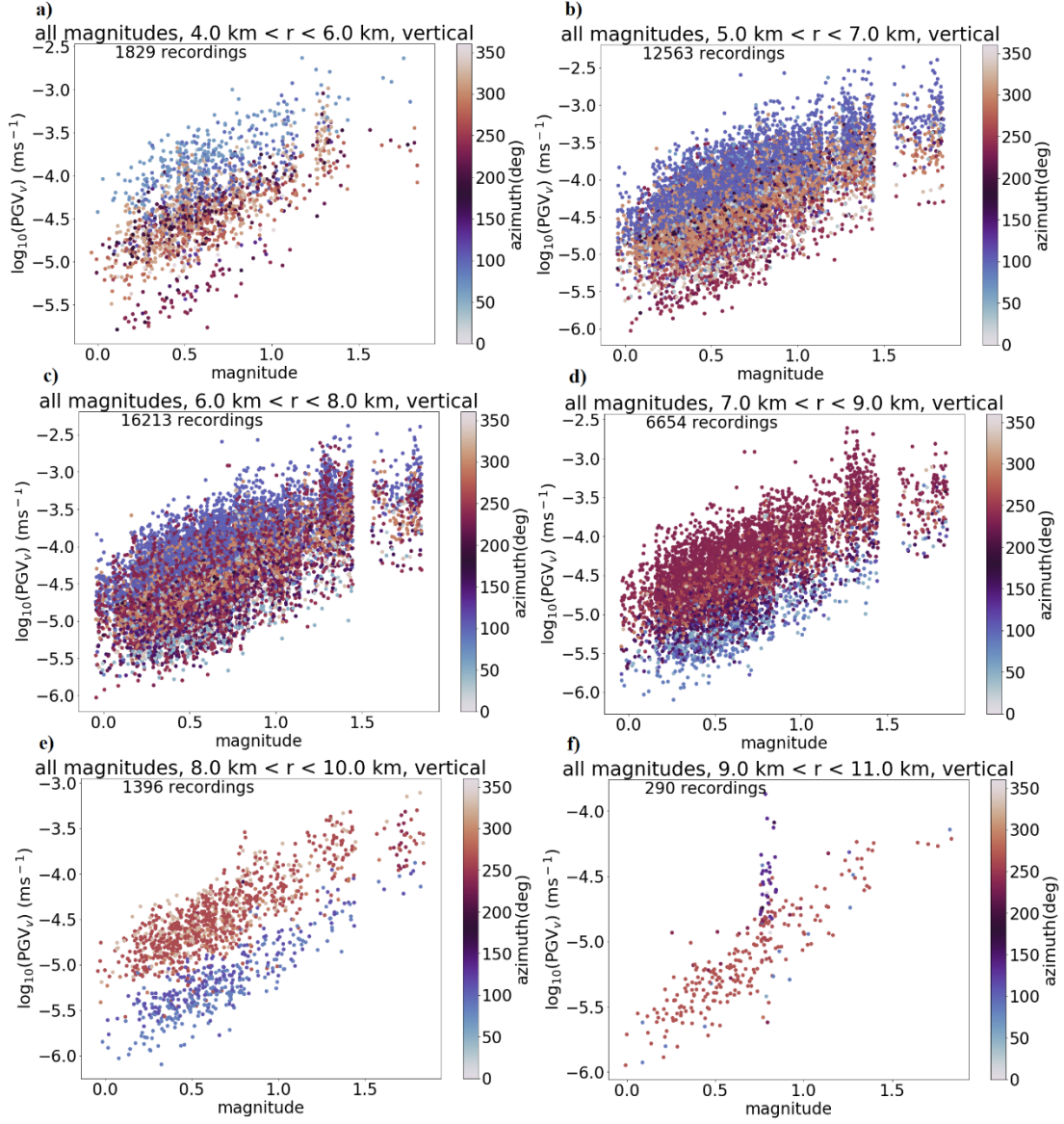


Figure 21. Vertical PGV variation with respect to magnitude (x-axis) and azimuth (colour) with 0° pointing directly north and with clockwise rotation.

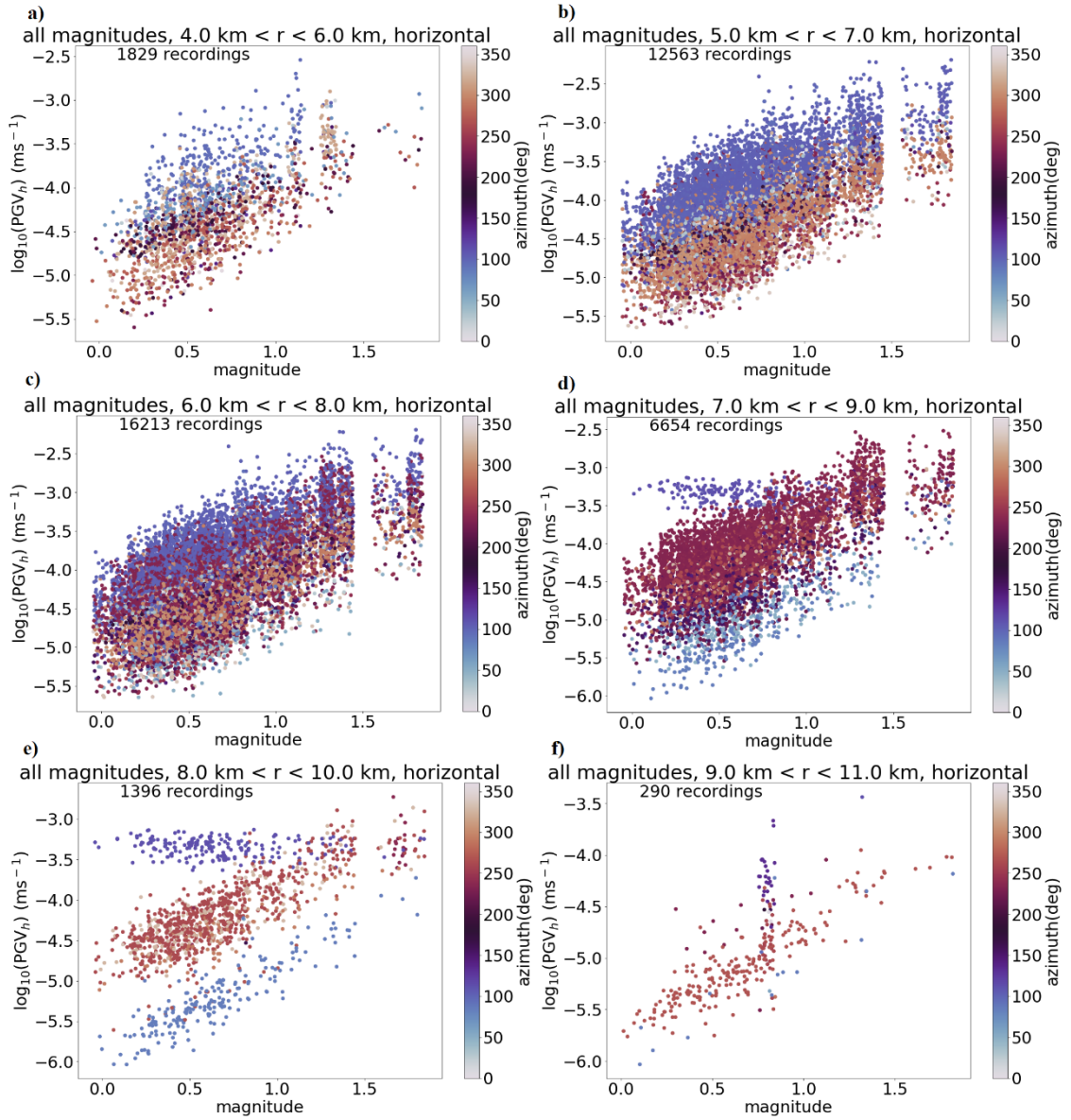


Figure 22. Horizontal PGV variation with respect to magnitude (x-axis) and azimuth (colour) with 0° pointing directly north and with clockwise rotation.

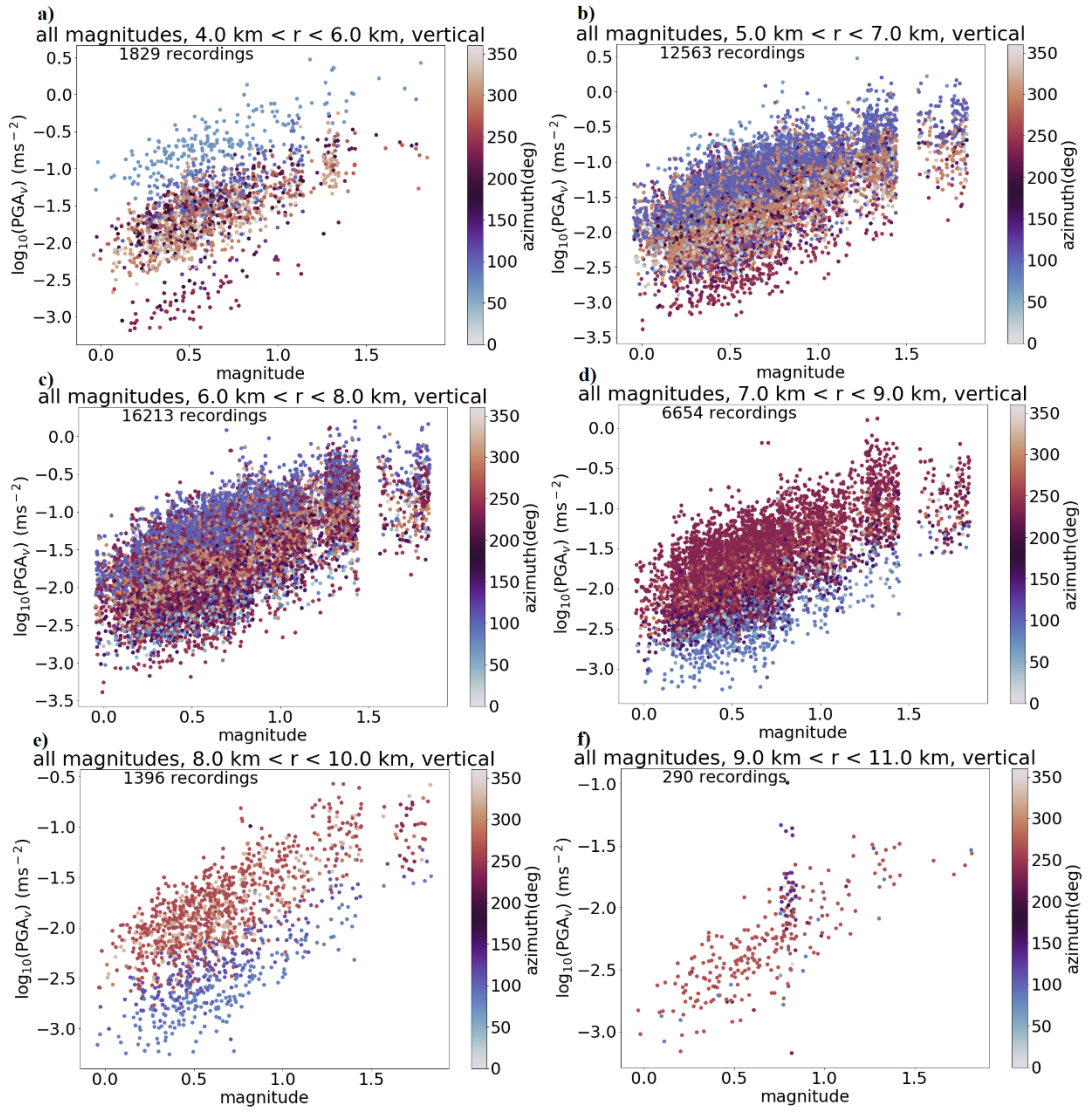


Figure 23. Vertical PGA variation with respect to magnitude (x-axis) and azimuth (colour) with 0° pointing directly north and with clockwise rotation.

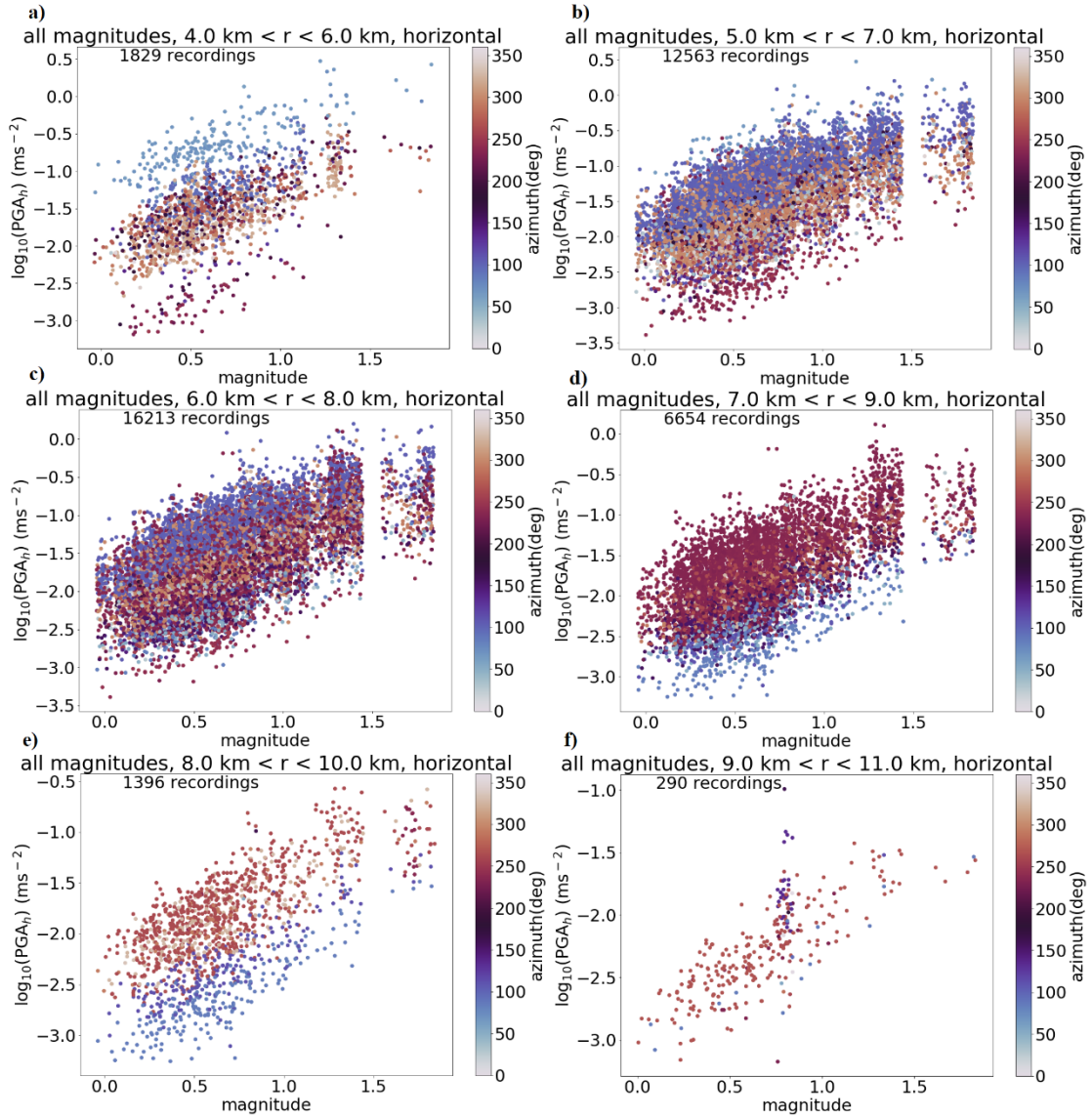


Figure 24. Horizontal PGA variation with respect to magnitude (x-axis) and azimuth (colour) with 0° pointing directly north and with clockwise rotation.

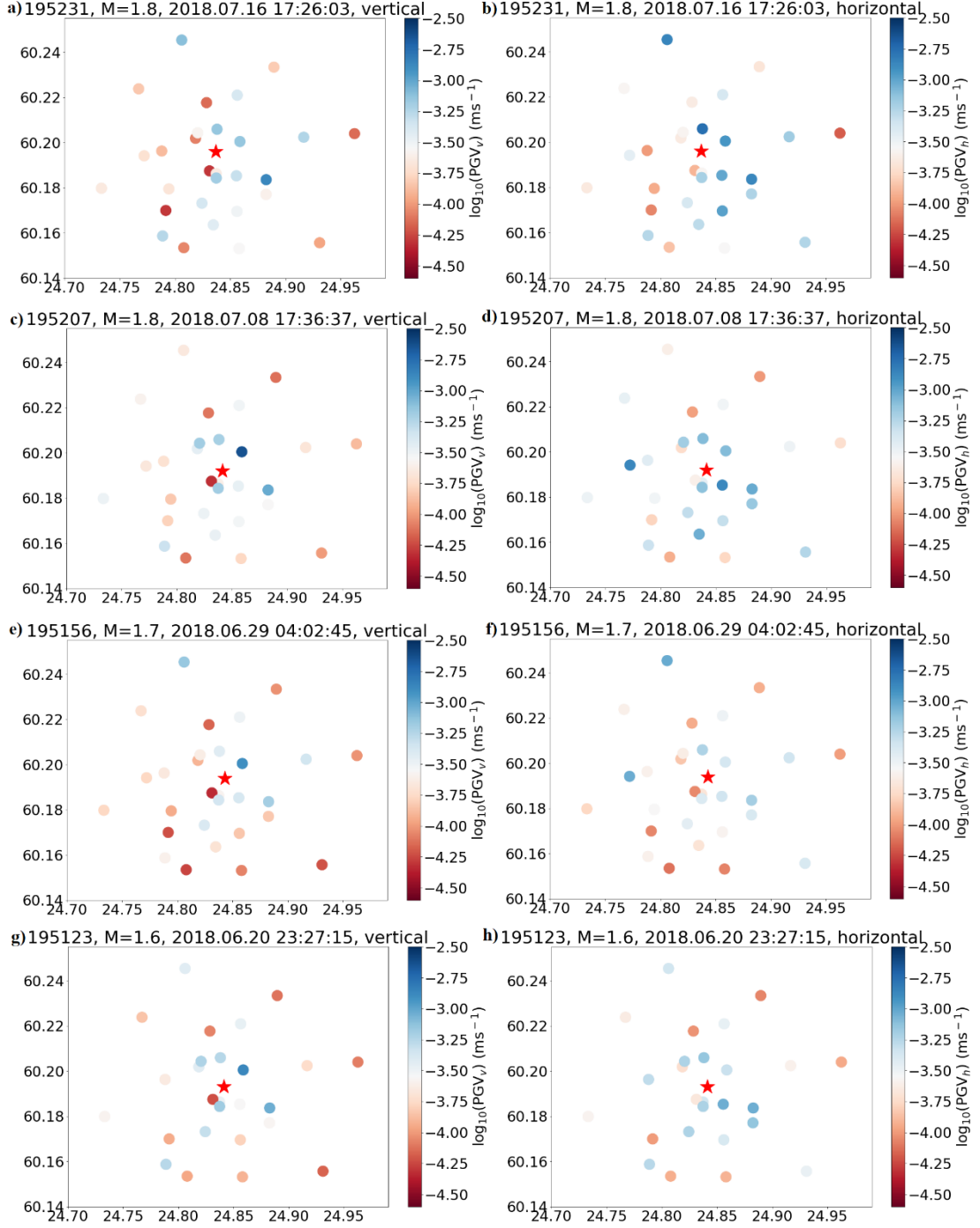


Figure 25. Examples of four events with their relative PGVs plotted around the epicentre. a) and b): $M_L = 1.8$ event on July 16th; c) and d): $M_L = 1.8$ event on July 8th; e) and f): $M_L = 1.7$ event on June 29th; g) and h): $M_L = 1.6$ event on June 20th. a), c), e) and g) are for vertical ground velocity on a common logarithm and b), d), f) and h) are for horizontal.

6. DISCUSSION

The results of the work are

1. The ground motion database consists of peak ground motions in vertical and horizontal direction for displacement, velocity, and acceleration, making it a useful item for studying peak ground motions at and around Otaniemi area.

2. The new model ON21 predicts vertical and horizontal peak ground velocity and peak ground acceleration. Of the four motion parameters, vertical peak ground velocity has the smallest standard deviation.

3. The azimuth of the measurements seem to affect the peak ground motion values, but not significantly and the effect seems to be strongest at very close distances to epicentre.

The results are discussed in more detail next.

6.1 The Peak Ground Motion Database

The first objective of this thesis was to compile a database of peak ground motion values. The CSV-file created and summarized in Table 3 provides an easy access to reliable peak ground motion values in both vertical and horizontal directions. Appendix 2 has an example of the CSV-file. The file contains more than 20,000 lines and 23 columns. The full dataset is available by request from the author. Every event from the dataset of 2018 stimulation is not included for two main reasons: event, id of which is 195197, caused an unexpected error in retrieving the data, and the events that had less than 2 km depth were omitted. The 204 events that are in the new database were enough for creating a new GMPE, as is discussed next.

6.2 Model ON21 Ground Motion Prediction Equation

Due to the available dataset, the model ON21 GMPE for Otaniemi is best used for predicting peak ground motions at the Finnish local magnitude range of 0.0 to 1.8 within distances of 0 km to 20 km. At magnitudes and distances outside this range its accuracy starts to suffer (Bommer et al. 2007). The 1σ -limits seem to hold in most if not all the data points at distances exceeding 10 km. Closer than 10 km, the PGV values deviate more, although this is also a result of having more data when the hypocentral distance is between 5 and 9 kilometres than when it is above 9 km. Initial tests of using any event that has the selected magnitude within the error limits of its magnitude would have increased the data points per individual magnitude, but would have also made the standard deviation much larger. This would have decreased the precision of the models, while not improving their accuracy. Thus, it was decided that only events with the same measured magnitude are used in the models. The horizontal ground motion models have larger absolute standard deviation than the vertical models, and the vertical PGA model has a larger absolute standard deviation than the vertical PGV model. Relative standard deviations cannot be compared due to the average values of the models being so close to zero. The greater standard deviations of the horizontal models compared to the vertical models may be due to there being more deviation in horizontal peak ground motion than in vertical. The greater standard deviation of the PGA models compared to PGV models may be due to the PGA values being larger than the PGV values.

Previous models for Otaniemi area, a model by Douglas et al. (2013) and Fin17 by ISUH according to Ader et al. (2020) underestimated ground motion at magnitudes above 1.2 before their correction (Arup 2018). Figure 20 shows the measured values from different stations with certain magnitudes. The ST1 satellite network (pink dots) on average gave smaller values for ground motion than the geophone stations and broadband seismometers of ISUH.

The decision to limit the effective range of the GMPE at 20 km also has effects on the shape of the model. Most localized GMPEs overpredict the peak ground motion past the distance they have received data from (Bommer et al. 2007). Quick testing indicated that

especially the horizontal PGA model tends to predict on average higher values at $M_L = 1.7$ and 1.8 . By comparing subfigures f in Figures 18 and 19, we can see that Figure 19f with the horizontal PGA has the cloud of data points at a slightly lower level compared to the model, than in figure 18f with the vertical PGA.

Because the model's functional form (Equation 3) has but two variables, and the error estimate σ is but the standard deviation of the measurements, the model does not take into account the epistemic uncertainties of the azimuth, the fault-mechanism and site-effects, and only considers the aleatory variability. The `curve_fit()`-function requires the errors of each measurement to better the fit. However, the dataset for the 2018 stimulation events did not contain error values for the depth estimates. As such, all depths were assumed to have a 100 m error, which is the accuracy at which the depths were stored in the original database. This was used with the horizontal error values to calculate a 3D error-vector for each event, which was then used as the sigma-input of the `curve_fit()`-function. Likely because the error was set large values, the covariance matrix and the standard deviation calculated from it became unrealistically high. In future, a better error-estimate for the depths could help decrease the error estimate by allowing it to be calculated directly from the covariance matrix received from the `curve_fit()`-function.

Appendix 1: Python Functions has some functions used in creating the GMPEs as well as a few functions that use the model ON21. These functions take in magnitude and distance to calculate the PGV or PGA value with the error limits, and functions that take in magnitude and PGV or PGA to calculate the distance at which these values are likely to occur, according to model ON21. Figure 26 has three examples of function "on21velvrplot" with a) an estimate of distance for 0.1 mms^{-1} vertical PGV for a $M_L = 1.0$ event, b) 1.0 mms^{-1} for a $M_L = 1.5$ event and c) 1.0 mms^{-1} for an $M_L = 1.8$ event. The 1σ -value of 0.598 translates roughly to a 4.487 km error in distance.

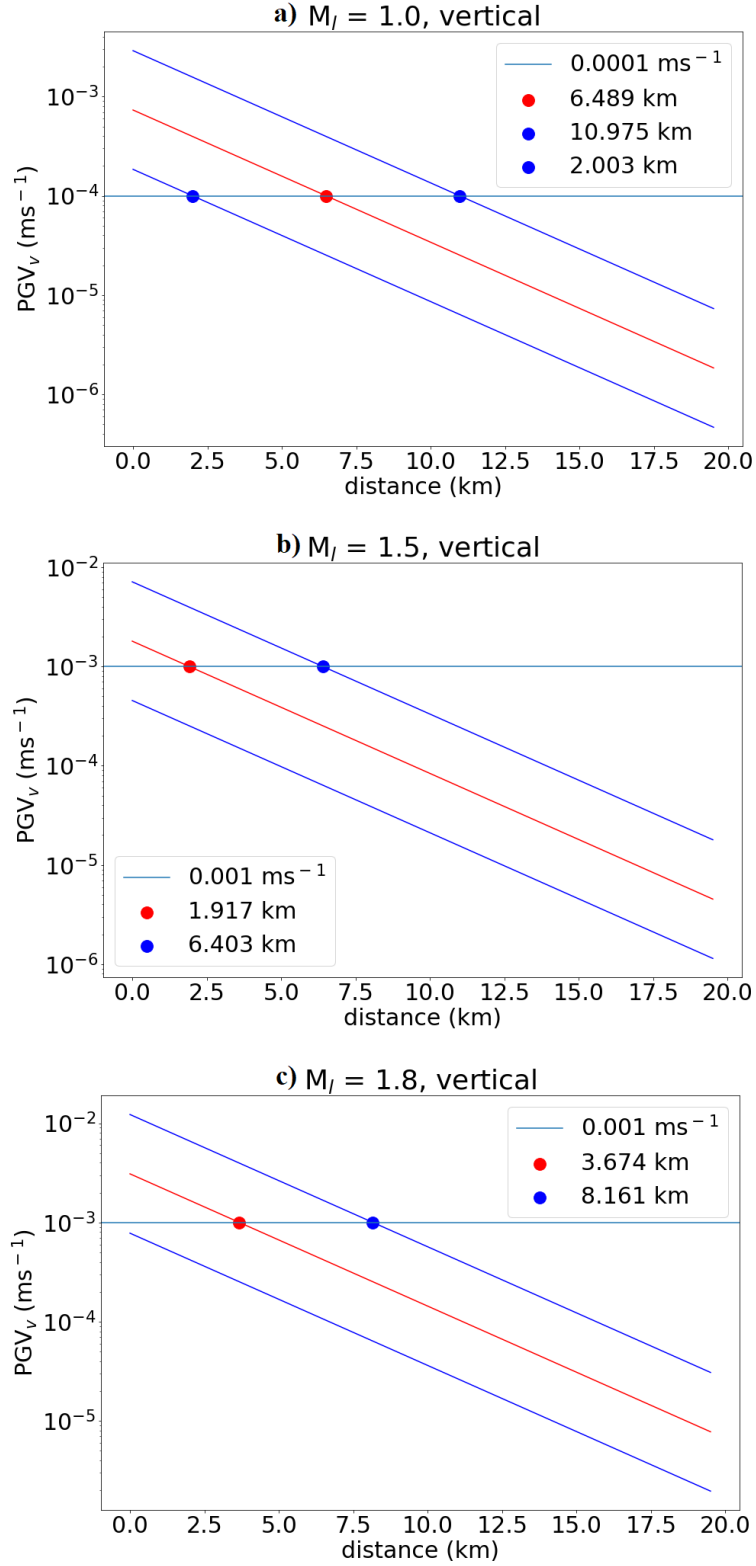


Figure 26. Examples of function “on21velvrplot” that is detailed in Appendix 1: Python Functions. a) an estimate of distance for 0.1 mms^{-1} vertical ground velocity for a $M_L = 1.0$ event, b) 1.0 mms^{-1} for a $M_L = 1.5$ event and c) 1.0 mms^{-1} for an $M_L = 1.8$ event. The 1σ -value of 0.598 translates to a 4.487 km error in distance.

A good test for the model ON21 is to plot the model with its 1σ values together with PGV measurements from the 2020 stimulation. While most events from the stimulation had negative magnitudes, proving a challenge for the range of ON21, a few also had positive magnitudes within the original database's limits (Veikkolainen et al. 2020). The new events could help calibrate the model further, and to test the limits of the model ON21 when magnitude is lower than the magnitudes of the model's reference events.

6.3 Effect of Azimuth

Figures 16, 17, 18 and 19 with the models also show the azimuths of the measurements. Especially at distances above 10 km, the residuals at 12 km with azimuth at around 200° to 250° seem to be positive in comparison to the measurements at around 11 km and 250° to 300° , and distances above 13 km and azimuths at around 50° to 100° . Figure 21, 22, 23 and 24 shows all the measurements from 2 km 'rings' around Otaniemi with their azimuths. The peak ground motion increases with increasing magnitude, but at distances between 4 km to 9 km the highest values come from azimuths at around 100° , with 300° at middle and lowest from around 250° . However, subfigures d in Figures 21, 22, 23 and 24 with distances at 7 km to 9 km and the consequent distances up to 11 km, show that the highest peak ground motion values come from around 225° and the values from around 100° become smaller in relation to the others. Figures 21, 22, 23 and 24 also show some curvature of the data; $M_L = 0.0$ seems to on average have higher values than $M_L = 0.1$. This is likely caused by there being less data for events of $M_L = 0.0$ than of $M_L = 0.1$. Similar effect at $M_L = 1.8$ seemingly having lower values than $M_L = 1.7$ and even $M_L = 1.4$ is most likely caused by the fact that there are less events at the higher magnitudes than there are for $M_L = 1.4$ and lower. Less events in the database tend to mean lesser variation in the measured velocities.

The dark blue clouds of measurements in Figures 22d and e are from an ST1 satellite station at Munkkisaari. It is likely these measures are simple outliers due to the station instrumentation itself and do not represent any real phenomenon. While the colours

denoting the azimuth seemingly give strong indication towards the azimuth's effect to the peak ground motion values, the rings' thickness of 2 km may be a stronger factor in some cases. Certain stations, as seen from Figure 5 have similar azimuths but very different distances. The clouds of colours in Figures 21, 22, 23 and 24 are thicker in some cases with thinner clouds under or over them, indicating an effect caused by the azimuth. However, the effect may be not as strong as the figures may make it seem.

Studies suggest that the energy radiation around the Otaniemi borehole is not uniform (e.g. Hillers et al. 2020, Leonhardt et al. 2021), and Douglas (2011) considers site-specific GMPEs more important at low magnitudes, in part because of the differences in local magnitude ranges. While Figures 21, 22, 23 and 24 indicate that the azimuth does play a role in the measured peak ground motion value, Figure 25, with events of $M_L = 1.8, 1.7$ and 1.6 does not show any clear azimuth effect. Comparing Figures 21, 22, 23 and 24 to Figure 25 could indicate that the azimuth does not have as big an effect as the hypocentral distance. According to Hillers et al. (2020), the energy from P-waves radiates most strongly to the north, while S-waves are the strongest to the east and west. In Figures 25 a, b, c and d this is not clearly seen. North-east of the borehole is the higher relative PGV in all figures, with south-west being lower. Figure 4 shows that the main lithological units and fault-structures have a north-east–south-west lineation, but Leonhardt et al. (2021) conclude that the slip causing the earthquakes underneath Otaniemi occurs mainly at a strike direction of north-north-west–south-south-east. The vertical PGV values around the strike line do seem to be affected by the azimuth, especially close to the epicentre, but for horizontal values the effect is not as clear.

7. CONCLUSIONS

The 2018 induced events from Otaniemi show a clear relation between the magnitude of the event, the hypocentral distance at which it was recorded, and its peak ground motion values. The created database of the peak ground motion values can be used individually or with other databases to solve for GMPEs at the Otaniemi region. While site-specific

GMPEs are important especially at low magnitudes, for areas similar to Otaniemi, the model ON21 can be used in initial PSHA. Similar areas to Otaniemi must be classified as being stable tectonic region, the earthquakes must occur at relatively shallow depths, and the environment has to be considered a hard-rock environment with similar expected body wave velocities.

The methodology for creating a GMPE presented in this work is robust enough to follow easily but acknowledges the need of individual testing and fitting. As such, the methods discussed in detail in chapter 4 can be used in future GMPE projects, with reference to this study.

The GMPE model ON21 is ready to be tested at Otaniemi region with more recent earthquakes. The 1σ -range is large enough to have most recorded events within its bounds, with largest deviation at distances less than 10 km, although this is likely due to the available data. Fault-mechanism, azimuth to the strike direction of the fault, and the epistemic uncertainties of site condition and the aleatory variability of peak ground motions due to earthquake processes are all included in the 1σ -value. This inflated the sigma, although it is still less than the 1σ -uncertainties for previous two GMPEs used in the same region. While the uncertainties are smaller, likely owing to the greater database of events, the model should be tested first with data from events that occurred during the 2020 stimulation and modified if found necessary.

There is evidence that the azimuth of the recording station to the strike-direction of the fault affects the peak ground motion values at distances less than 10 km. Due to the lack of data it is unclear how quickly this effect diminishes as distance grows. In future a mathematical modelling of the azimuth's effect could be experimented to further improve the model ON21.

8. ACKNOWLEDGEMENTS

I would like to thank my thesis supervisors M.Sc. Tommi Vuorinen and Ph.D. Päivi Mäntyniemi for the help, tips, and teaching they provided during the completion of this thesis. I thank Ph.D. Ludovic Fülöp for his helpful comments regarding the filter-ranges for magnitudes and azimuths when creating the figures, and Ph.D. Gregor Hillers for his help and suggestions with the programming, and Ph.D. student Annukka Rintamäki for providing information about the fault-structures. I would also like to thank M.Sc. student Ilkka Kettunen for his help with QGIS-program, and M.Sc. Riina Aapasuo for her bachelor's study, which helped me greatly in understanding the theory behind instrument response removal.

9. REFERENCES

- Ader, T., Chendorain, M., Free, M., Saarno, T., Heikkinen, P., Malin, P.E., Leary, P., Kwiatak, G., Dresen, G., Bluemle, F. & Vuorinen, T.A.T. 2020, Design and implementation of a traffic light system for deep geothermal well stimulation in Finland, *Journal of Seismology* 24, 991–1014v <https://doi-org.libproxy.helsinki.fi/10.1007/s10950-019-09853-y>
- Anderson, J.G. 2003, Strong-motion seismology, *International Geophysics Series*, **81**(B), p. 937-966, <https://doi-org.libproxy.helsinki.fi/10.1002/rog.1991.29.s2.700>
- Arup, 2018, Implementation of Traffic Light System for Seismic Monitoring during OTN-III Well Stimulation: Summary of Findings, report, Otaniemi Geothermal Doublet, St1 Deep Heat, p. 35
- Asch, G. 2009, Seismic Recording Systems, in *New Manual of Seismological Observatory Practice 2 (NMSOP 2)*, edit. P. Bormann, Deutsches GeoForschungsZentrum GFZ, Potsdam, p. 20.
- Atkinson, G.M. 2008, Ground-Motion Prediction Equations for Eastern North America from a Referenced Empirical Approach: Implications for Epistemic Uncertainty, *Bulletin of the Seismological Society of America*, **98**(3), p. 1304-1318. DOI 10.1785/0120070199.
- Bendat, J.S. & Piersol, A.G. 2011, *Random Data: Analysis and Measurement Procedures*, 4. p., Wiley, Somerset, p. 1-55
- Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y. & Wassermann, J. 2010, ObsPy; a python toolbox for seismology, *Seismological Research Letters*, **81**(3), p. 530-533. DOI 10.1785/gssrl.81.3.530.
- Bolt, B.A. & Abrahamson, N.A. 2003, Estimation of Strong Seismic Ground Motions, *teoksessa International Handbook of Earthquake & Engineering Seismology*, p. 983-1002.
- Bommer, J.J., Stafford, P.J., Alarcon, J.E. & Akkar, S. 2007, The Influence of Magnitude Range on Empirical Ground-Motion Prediction, *Bulletin of the Seismological Society of America*, **97**(6), p. 2152-2170. DOI 10.1785/0120070081.
- Bormann, P. & Wielandt, E. 2013, Seismic Signals and Noise, in *International Handbook of Earthquake & Engineering Seismology*, edit. P. Bormann, Deutsches GeoForschungsZentrum GFZ, Potsdam, s. 62.
- Bormann, P., Klinge, K. & Wendt, S. 2014, Data Analysis and Seismogram Interpretation, in *New Manual of Seismological Observatory Practice 2 (NMSOP 2)*, edit. P. Bormann, Deutsches GeoForschungsZentrum GFZ, Potsdam.
- Buijze, L., van Bijsterveldt, L., Cremer, H., Paap, B., Veldkamp, H., Wassing, B.B.T., van Wees, J., van Yperen, Guido C. N, ter Heege, J. H. & Jaarsma, B. 2019, Review of induced seismicity in geothermal systems worldwide and implications

- for geothermal systems in the Netherlands, *Geologie en Mijnbouw; Netherlands Journal of Geosciences*, **98**. DOI 10.1017/njg.2019.6.
- Campbell, K.W. 1981, Near-source attenuation of peak horizontal acceleration, *Bulletin of the Seismological Society of America*, **71**(6), p. 2039-2070.
- Campbell, K.W. 2003, Strong-Motion Attenuation Relations, in *International Handbook of Earthquake & Engineering Seismology*, edit. W.H.K. Lee, H. Kanamori, P.C. Jennings & C. Kisslinger, p. 1003-1012.
- Campbell, K.W. 2010, Ground Motion Simulation Using the Hybrid Empirical Method: Issues and Insights, in *Earthquake Data in Engineering Seismology*. [Online]. Dordrecht: Springer Netherlands. pp. 81–95.
- Douglas, J. 2011, Investigating Possible Regional Dependence in Strong Ground Motions, in *Earthquake Data in Engineering Seismology*. [Online]. Dordrecht: Springer Netherlands. pp. 29–38.
- Douglas, J. & Aochi, H. 2008, A Survey of Techniques for Predicting Earthquake Ground Motions for Engineering Purposes, *Surveys in Geophysics*, **29**(3), p. 187-220. DOI 10.1007/s10712-008-9046-y.
- Douglas, J., Edwards, B., Convertito, V., Sharma, N., Tramelli, A., Kraaijpoel, D., Cabrera, B.M., Maercklin, N. & Troise, C. 2013, Predicting ground motion from induced earthquakes in geothermal areas, *Bulletin of the Seismological Society of America*, **103**(3), p. 1875-1897. DOI 10.1785/0120120197.
- Fülöp, L., Jussila, V., Aapasuo, R., Vuorinen, T.A.T., Mäntyniemi, P.B., Hämäläinen, J. & Suolanen, V. 2019, Evolving the Fennoscandian GMPEs (EVOGY), SAFIR2018–The Finnish Research Programme on Nuclear Power Plant Safety 2015-2018: Final report, , p. 422-433.
- Gadallah, M.R. & Fisher, R. 2009, Seismic Data Processing, in *Exploration Geophysics*, Springer, Berlin Heidelberg, p. 85-148.
- Geological Survey of Finland (GTK), 2017, Bedrock of Finland 1:200 000. Retrieved 5.4.2021. https://hakku.gtk.fi/locations/search?location_id=2
- Giardini, D. 2009, Geothermal quake risks must be faced, *Nature (London)*; *Nature*, **462**(7275), p. 848-849. DOI 10.1038/462848a.
- Hansen, P.C., Pereyra, V. & Scherer, G. 2013, Least squares data fitting with applications, Johns Hopkins University Press, Baltimore, Md, p. 1-46
- Hillers, G., Vuorinen, T.A.T., Uski, M.R., Kortström, J.T., Mäntyniemi, P.B., Tiira, T., Malin, P.E. & Saarno, T. 2020, The 2018 Geothermal Reservoir Stimulation in Espoo/Helsinki, Southern Finland: Seismic Network Anatomy and Data Features, *Seismological Research Letters*, **91**(2), p. 770-786. DOI 10.1785/0220190253.
- Kafaei Mohammadnejad, A., Mousavi, S.M., Torabi, M., Mousavi, M. & Alavi, A.H. 2012, Robust attenuation relations for peak time-domain parameters of strong ground motions, *Environmental Earth Sciences*, **67**(1), p. 53-70. DOI 10.1007/s12665-011-1479-9.

- Kawase, H. 2003, Site effects on strong ground motions, *International Geophysics Series*, **81(B)**, p. 1013-1030.
- Kukkonen, I.T., Kuusisto, M., Lehtonen, M. & Peltonen, P. 2008, Delamination of eclogitized lower crust: Control on the crust–mantle boundary in the central Fennoscandian shield, *Tectonophysics*, **457**(3), p. 111-127. DOI 10.1016/j.tecto.2008.04.029.
- Kukkonen, I.T. & Lauri, L.S. 2009, Modelling the thermal evolution of a collisional Precambrian orogen: High heat production migmatitic granites of southern Finland, *Precambrian Research*, **168**(3), p. 233-246. DOI 10.1016/j.precamres.2008.10.004.
- Kwiatek, G., Saarno, T., Ader, T., Bluemle, F., Bohnhoff, M., Chendorain, M., Dresen, G., Heikkinen, P., Kukkonen, I.T., Leary, P., Leonhardt, M., Malin, P., Martínez-Garzón, P., Passmore, K., Passmore, P., Valenzuela, S. & Wollin, C. 2019, Controlling fluid-induced seismicity during a 6.1-km-deep geothermal stimulation in Finland.
- Lekshmy, P.R. & Raghukanth, S.T.G. 2021, A hybrid genetic algorithm-neural network model for power spectral density compatible ground motion prediction, *Soil dynamics and earthquake engineering* (1984), **142**, p. 106528. DOI 10.1016/j.soildyn.2020.106528.
- Leonhardt, M., Kwiatek, G., Martinez-Garzon, P., Bohnhoff, M., Saarno, T., Heikkinen, P. & Dresen, G. 2021, Seismicity during and after stimulation of a 6.1 km deep enhanced geothermal system in Helsinki, Finland, *Solid earth (Göttingen)*, **12**(3), p. 581-594. DOI 10.5194/se-12-581-2021.
- Moré, J.J., Garbow, B.S. & Hillstom, K.E. 1980, User Guide for MINPACK-1, User guide for MINPACK-1, p. 9.
- Pezeshk, S., Zandieh, A. & Tavakoli, B. 2011, Hybrid Empirical Ground-Motion Prediction Equations for Eastern North America Using NGA Models and Updated Seismological Parameters, *Bulletin of the Seismological Society of America*, **101**(4), p. 1859-1870. DOI 10.1785/0120100144.
- Råde, L. & Westergren, B. 2004, *Mathematics handbook for science and engineering*, 5. p., Studentlitteratur, Lund, p. 311-321
- Scherbaum, F. 2001, Of poles and zeros: fundamentals of digital seismology, 2. p., Kluwer Academic Publishers, Dordrecht, p. 12-38, 48-65, 95-116
- Schultz, R., Beroza, G., Ellsworth, W. & Baker, J. 2020, Risk-Informed Recommendations for Managing Hydraulic Fracturing–Induced Seismicity via Traffic Light Protocols, *Bulletin of the Seismological Society of America*. DOI 10.1785/0120200016.
- Schön, J. 2011, *Physical properties of rocks: a workbook*, Elsevier, Oxford, p. 154-161.
- scipy.optimize.curve_fit — SciPy v1.6.2 Reference Guide 2021, (Retrieved 26.4.2021).
- Shakal, A.F., Huang, M.J. & Graizer, V.M. 2003, Strong-Motion Data Processing, in *International Handbook of Earthquake & Engineering Seismology*, edit. W.H. Lee, H. Kanamori, P.C. Jennings & C. Kisslinger, p. 967-982.

- Smith, W.A. & Randall, R.B. 2016, Cepstrum-based operational modal analysis revisited: A discussion on pole–zero models and the regeneration of frequency response functions, *Mechanical systems and signal processing*, **79**, p. 30-46. DOI 10.1016/j.ymssp.2016.02.030.
- Veikkolainen, T. & Kukkonen, I.T. 2019, Highly varying radiogenic heat production in Finland, Fennoscandian Shield, *Tectonophysics*, **750**, p. 93-116. DOI 10.1016/j.tecto.2018.11.006.
- Veikkolainen, T., Oinonen, K., Vuorinen, T.A.T, Kortström, J.T., Mäntyniemi, P.B., Lindblom, P., Uski, M.R. & Tiira, T. 2021, Helsingin seisminen asemaverkko ja seismisyys 2020, Seismologian instituutti, Geotieteiden ja maantieteen osasto, Helsingin yliopisto.
- Veikkolainen, T., Uski, M.R., Voutilainen, A., Vuorinen, T.A.T, Oinonen, K., Kortström, J.T., Luhta, T., Mäntyniemi, P.B. & Tiira, T. 2020, Seismisyys Espoon Otaniemen syväreikähankkeen viimeistelyvaiheessa, Seismologian instituutti, Geotieteiden ja maantieteen osasto, Helsingin yliopisto.
- Vuorinen, T.A.T., Tiira, T., Uski, M.R. & Lund, B. 2018, Updated Fennoscandian GMPE. University of Helsinki, Institute of Seismology, p. 73.
- Wielandt, E. 2012, Seismic Sensors and their calibration, in *New Manual of Seismological Observatory Practice 2 (NMSOP 2)*, edit. P. Bormann, Deutsches GeoForschungsZentrum GFZ, Potsdam, s. 1-51.

APPENDIX 1: PYTHON FUNCTIONS

```

import numpy as np
import matplotlib.pyplot as plt
from math import degrees
from math import atan

def campbell(stuff,c1,c2,c3):
    #A function for Campbell's (2003) model of GMPE to be used with scipy.optimize.curve_fit
    #'stuff' is a two-part list: stuff[0] should have the event distances, and stuff[1] the event
    magnitudes
    filtd = stuff[0]
    M = stuff[1]
    C = c1+c2*M-c3*filtd
    return C

def azimuth(lat_event, lon_event, lat_station, lon_station):
    #function for receiving the azimuth between north (0 deg) and station (clockwise)
    lat_event = float(lat_event.val)
    lon_event = float(lon_event.val)
    lat_station = float(lat_station)
    lon_station = float(lon_station)
    azim = float(degrees(atan((lon_station-lon_event)/(lat_station-lat_event))))
    if((lon_station-lon_event<0) & (lat_station-lat_event<0)):
        azim+=180
    elif((lon_station-lon_event<0) & (lat_station-lat_event>=0)):
        azim=360+azim
    elif((lon_station-lon_event>=0) & (lat_station-lat_event<0)):
        azim=180+azim

    return int(round(azim))

#=====Fourier domain differentials
def derivate_in_F(trace):
    #derivates a trace in frequency domain
    #returns the derivated trace in time-domain and its maximum value
    #velocity derivated gives acceleration
    value = np.fft.rfft(trace.data)
    n = len(trace.data)
    sampling_rate = trace.stats.sampling_rate
    freqs = np.fft.rfftfreq(n, 1./sampling_rate)

    inv_der = value*freqs*2*np.pi*1j
    inv_der = np.fft.irfft(inv_der)

    max = np.max(np.abs(inv_der))
    return inv_der, max

```

```

def integrate_in_F(trace):
    #integrates a trace in frequency domain
    #returns the integrated trace in time-domain and its maximum value
    #velocity integrated gives displacement
    value = np.fft.rfft(trace.data)
    n = len(trace.data)
    sampling_rate = trace.stats.sampling_rate
    freqs = np.fft.rfftfreq(n, 1./sampling_rate)

    inv_int = (value[1:]/freqs[1:])*(1j/(2*np.pi))
    inv_int[0] = 0
    inv_int = np.fft.irfft(inv_int)

    max = np.max(np.abs(inv_int))
    return inv_int, max

#=====PGV
def on21velv(m,r):
    #Returns PGV in m/s and deviation when given magnitude m and distance r (km)
    vmed = 10**(-3.916+0.781*m-0.1333*r)
    vdevupper = 10**(-3.916+0.781*m-0.1333*r+0.598)
    vdevlower = 10**(-3.916+0.781*m-0.1333*r-0.598)
    return vmed, vdevupper, vdevlower

def on21velh(m,r):
    #same as on21velv but for horizontal movement
    vmed = 10**(-3.925+0.786-0.137*r)
    vdevupper = 10**(-3.925+0.786-0.137*r+0.676)
    vdevlower = 10**(-3.925+0.786-0.137*r-0.676)
    return vmed, vdevupper, vdevlower

def on21velvr(m, pgv):
    #Returns distance (km) and deviation when given magnitude m and peak ground velocity
    pgv (m/s)
    vmed = (-3.916+0.781*m-np.log10(pgv))/0.1333
    vdevupper = (-3.916+0.781*m-np.log10(pgv)+0.598)/0.1333
    vdevlower = (-3.916+0.781*m-np.log10(pgv)-0.598)/0.1333
    return vmed, vdevupper, vdevlower

def on21velhr(m,pgv):
    #same as on21velvr but for horizontal movement
    vmed = (-3.925+0.786*m-np.log10(pgv))/0.137
    vdevupper = (-3.925+0.786*m-np.log10(pgv)+0.676)/0.137
    vdevlower = (-3.925+0.786*m-np.log10(pgv)-0.676)/0.137
    return vmed, vdevupper, vdevlower

def on21velvrplot(m, pgv, log=True):

```

```

#Draws the vertical GMPE and returns distances where the desired PGV (m/s) occurs
#m=local magnitude
plt.figure(figsize=(15,10))
plt.rcParams.update({'font.size': 30})
dist=np.arange(0,20,0.5)

on21pgv=on21velv(m,dist)
plt.plot(dist,on21pgv[0], c='r')

plt.plot(dist,on21pgv[1], c='b')
plt.plot(dist, on21pgv[2], c='b')

r=on21velvr(m,pgv)
plt.scatter(r[0],pgv, s=250, c='r', label=('%.2f km' %r[0]))
plt.scatter(r[1],pgv, s=250, c='b', label=('%.2f km' %r[1]))
if(r[2]>0):
    plt.scatter(r[2],pgv, s=250, c='b', label=('%.2f km' %r[2]))

plt.axhline(pgv, label=('%.g ms$^-$ $^1$' %pgv))
plt.xlabel('distance (km)')
plt.ylabel('PGV$$_v$ (ms$^-$ $^1$)')
if(log==True):
    plt.yscale('log')
plt.title('M$_l$ = %.1f, vertical' %m)
plt.legend()

return r[0], r[1]

def on21velhrplot(m, pgv, log=True):
    #Draws the horizontal GMPE and returns distances where the desired PGV (m/s) occurs
    #m=local magnitude
    plt.figure(figsize=(15,10))
    plt.rcParams.update({'font.size': 30})
    dist=np.arange(0,20,0.5)

    on21pgv=on21velh(m,dist)
    plt.plot(dist,on21pgv[0], c='r')

    plt.plot(dist,on21pgv[1], c='b')
    plt.plot(dist, on21pgv[2], c='b')

    r=on21velhr(m,pgv)
    plt.scatter(r[0],pgv, s=250, c='r', label=('%.2f km' %r[0]))
    plt.scatter(r[1],pgv, s=250, c='b', label=('%.2f km' %r[1]))
    if(r[2]>0):
        plt.scatter(r[2],pgv, s=250, c='b', label=('%.2f km' %r[2]))

    plt.axhline(pgv, label=('%.g ms$^-$ $^1$' %pgv))
    plt.xlabel('distance (km)')
    plt.ylabel('PGV$$_h$ (ms$^-$ $^1$)')
    if(log=True):

```

```

    plt.yscale('log')
    plt.title('ML = %1.1f, horizontal' %m)
    plt.legend()

    return r[0], r[1]

#=====PGA

def on21accv(m,r):
    #Returns PGA in m/s2 and deviation when given magnitude m and distance r (km)
    amed = 10**(-1.099+0.836*m-0.153*r)
    adevupper = 10**(-1.099+0.836*m-0.153*r+0.611)
    adevlower = 10**(-1.099+0.836*m-0.153*r-0.611)
    return amed, adevupper, adevlower

def on21acch(m,r):
    #same as on21accv but for horizontal movement
    amed = 10**(-1.235+0.991*m-0.153*r)
    adevupper = 10**(-1.235+0.991*m-0.153*r+0.642)
    adevlower = 10**(-1.235+0.991*m-0.153*r-0.642)
    return amed, adevupper, adevlower

def on21accvr(m, pga):
    #Returns distance (km) and deviation when given magnitude m and peak ground velocity
    pgv (m/s)
    amed = (-1.099+0.836*m-np.log10(pga))/0.153
    adevupper = (-1.099+0.836*m-np.log10(pga)+0.611)/0.153
    adevlower = (-1.099+0.836*m-np.log10(pga)-0.611)/0.153
    return amed, adevupper, adevlower

def on21acchr(m,pga):
    #same as on21velvr but for horizontal movement
    amed = (-1.235+0.991*m-np.log10(pga))/0.153
    adevupper = (-1.235+0.991*m-np.log10(pga)+0.642)/0.153
    adevlower = (-1.235+0.991*m-np.log10(pga)-0.642)/0.153
    return amed, adevupper, adevlower

def on21accvrplot(m, pga, log=True):
    #Draws the vertical motion GMPE and returns distances where the desired PGA (m/s2)
    occurs
    #m=local magnitude
    plt.figure(figsize=(15,10))
    plt.rcParams.update({'font.size': 30})
    dist=np.arange(0,20,0.5)

    on21pga=on21accv(m,dist)
    plt.plot(dist,on21pga[0], c='r')

    plt.plot(dist,on21pga[1], c='b')

```

```

plt.plot(dist, on21pga[2], c='b')

r=on21accvr(m,pga)
plt.scatter(r[0],pga, s=250, c='r', label=('%2.3f km' %r[0]))
plt.scatter(r[1],pga, s=250, c='b', label=('%2.3f km' %r[1]))
if(r[2]>0):
    plt.scatter(r[2],pga, s=250, c='b', label=('%2.3f km' %r[2]))

plt.axhline(pga, label=('%g ms$^-$^2$' %pga))
plt.xlabel('distance (km)')
plt.ylabel('PGA$_v$ (ms$^-$^2$)')
if(log==True):
    plt.yscale('log')
plt.title('M$_l$ = %1.1f, vertical' %m)
plt.legend()

return r[0], r[1]

```

```

def on21acchrplot(m, pga, log=True):
    #Draws the horizontal motion GMPE and returns distances where the desired PGA (m/s2)
    occurs
    #m=local magnitude
    plt.figure(figsize=(15,10))
    plt.rcParams.update({'font.size': 30})
    dist=np.arange(0,20,0.5)

    on21pga=on21acch(m,dist)
    plt.plot(dist,on21pga[0], c='r')

    plt.plot(dist,on21pga[1], c='b')
    plt.plot(dist, on21pga[2], c='b')

    r=on21acchr(m,pga)
    plt.scatter(r[0],pga, s=250, c='r', label=('%2.3f km' %r[0]))
    plt.scatter(r[1],pga, s=250, c='b', label=('%2.3f km' %r[1]))
    if(r[2]>0):
        plt.scatter(r[2],pga, s=250, c='b', label=('%2.3f km' %r[2]))

    plt.axhline(pga, label=('%g ms$^-$^2$' %pga))
    plt.xlabel('distance (km)')
    plt.ylabel('PGA$_h$ (ms$^-$^2$)')
    if(log==True):
        plt.yscale('log')
    plt.title('M$_l$ = %1.1f, horizontal' %m)
    plt.legend()

    return r[0], r[1]

```

APPENDIX 2: EXAMPLE OF ST1-18_DATABASE_PGM.CSV

	sta			M_	dep	depth_	fixed		eve	lat_e	eve	lon_e					azimu		PGA_ho				
id	n	network	M	err	th(error(_dep	th	distanc	nt_l	rror(nt_l	rror(PGD(mm	PGV(mm	PGA(m	filterin	th(de	PGD_hor	PGV_ho	r(mm/s	statl	statl
19	50	TL	Toppelun	0	.	530			6857.5					0.000115	0.04020	33.457	['highp		0.000163	0.12856		60.1	24.7
76	16	d_cubes	6	0.4	0.0	0.0	TOSI	21	9	14.0	34	15.0	06	6	38	5.0]	235	38	5	381882	14	14	
19	50	EV	Elfvik_cu	0	.	530			5619.5					0.000198		59.921	['highp		0.000280	0.06323		60.2	24.8
76	04	bes	6	0.4	0.0	0.0	TOSI	68	9	14.0	34	15.0	04	9	5703774	5.0]	318	85	79	631093	63	55	
19	50	EV	Elfvik_cu	0	.	530			5624.3					0.000132	0.04425	25.461	['highp		0.000187	0.01689		60.2	24.8
76	03	bes	6	0.4	0.0	0.0	TOSI	65	9	14.0	34	15.0	06	5	67	5.0]	318	76	06	889065	53	17	
19	50	EV	Elfvik_cu	0	.	530			5631.5					0.03625	20.716	['highp		0.000128	0.03630		60.2	24.8	
76	01	bes	6	0.4	0.0	0.0	TOSI	11	9	14.0	34	15.0	6865e-05	5	05	5.0]	316	16	06	468558	78	37	
19	50	M	ST1_satel	0	.	530			4479.3					0.000115		195.28	['highp		0.000162	0.08057			
76	A	K	6	0.4	0.0	0.0	TOSI	15	9	14.0	34	15.0	53	3791618	67	5.0]	67	39	6	207021	005	588	
19	50	EV	Elfvik_cu	0	.	530			5628.7					7.943419	0.02732	18.771	['highp		0.000112	0.04067		60.2	24.8
76	00	bes	6	0.4	0.0	0.0	TOSI	62	9	14.0	34	15.0	05	5	65	5.0]	315	12	43	187936	54	99	
19	50	EV	Elfvik_cu	0	.	530			5633.0					0.000120	0.04877	19.335	['highp		0.000170	0.03933		60.2	24.8
76	02	bes	6	0.4	0.0	0.0	TOSI	49	9	14.0	34	15.0	51	2	74	5.0]	317	6	14	046931	67	29	
19	50	HE		0	.	530			5649.0					7.356707	0.03265	17.958	['highp		0.000104	0.05532			
76	L3	HEL	6	0.4	0.0	0.0	TOSI	67	9	14.0	34	15.0	05	5	54	5.0]	210	38	04	806455	732	244	
19	50	WE	Single_cu	0	.	530			5858.9					4.001499		4.6757	['highp		5.658974	0.01167			
76	G	bes	6	0.4	0.0	0.0	TOSI	51	9	14.0	34	15.0	05	2657721	51	5.0]	255	05	62	467141	7954	9444	
19	50		Single_cu	0	.	530			5902.7					4.181481	0.00990	5.4661	['highp		5.913507	0.02235			
76	LTS	bes	6	0.4	0.0	0.0	TOSI	02	9	14.0	34	15.0	05	98	94	5.0]	133	05	48	352498	6958	5612	
19	50	TA	ST1_satel	0	.	530			4918.0					2.265506		23.198	['highp		3.203909	0.02571			
76	GC	K	6	0.4	0.0	0.0	TOSI	18	9	14.0	34	15.0	05	5007182	59	5.0]	278	05	1	494166	963	88	

19			0					5973.8					0.000165	0.02351	16.194	['high	0.000233	0.03215				
50	SS	Seurasaa	.		530			60447	60.1		24.8		3024701	9830260	87022	ass',	7729952	450806	22.9030	60.1	24.8	
76	10	ri_cubes	6	0.4	0.0	0.0	TOSI	1	9	14.0	34	15.0	9	8	14	5.0]	98	37	76	05108	8357	8181
19	SS		0					5974.9					9.448279	0.04372		['high	0.000133	0.09760				
50	00	Seurasaa	.		530			46388	60.1		24.8		56524e-	9370964	28.279	ass',	6188510	698763	39.9934	60.1	24.8	
76	B	ri_cubes	6	0.4	0.0	0.0	TOSI	13	9	14.0	34	15.0	05	5	64137	5.0]	97	23	3	523646	8398	8207
19			0					5980.0					4.492560	0.02489	18.389	['high	6.353440	0.06081				
50	SS	Seurasaa	.		530			39152	60.1		24.8		57464e-	1827258	48816	ass',	09443e-	539803	26.0066	60.1	24.8	
76	11	ri_cubes	6	0.4	0.0	0.0	TOSI	03	9	14.0	34	15.0	05	8	78	5.0]	98	05	5	63572	8364	821
19			0					5988.5					4.700647	0.01389	7.9669	['high	6.647718	0.03081				
50	SS	Seurasaa	.		530			47618	60.1		24.8		14737e-	9744204	39055	ass',	94774e-	312018	11.2669	60.1	24.8	
76	07	ri_cubes	6	0.4	0.0	0.0	TOSI	1	9	14.0	34	15.0	05	4	84	5.0]	97	05	81	532634	8384	8255
19			0					5992.7					7.667088	0.03894	23.976	['high	0.000108	0.04647				
50	SS	Seurasaa	.		530			35556	60.1		24.8		30528e-	4558683	13037	ass',	4290026	186548	33.9073	60.1	24.8	
76	12	ri_cubes	6	0.4	0.0	0.0	TOSI	81	9	14.0	34	15.0	05	6	93	5.0]	97	52	99	687557	8365	8262
19			0					5993.6					0.000195	0.08068	38.281	['high	0.000277	0.21920				
50	SS	Seurasaa	.		530			27835	60.1		24.8		9060358	7225845	10169	ass',	0529728	550219	54.1376	60.1	24.8	
76	17	ri_cubes	6	0.4	0.0	0.0	TOSI	14	9	14.0	34	15.0	4	7	36	5.0]	98	35	3	531977	8338	8251
19			0					5995.4					0.000136	0.05144	26.491	['high	0.000193	0.05678				
50	SS	Seurasaa	.		530			78169	60.1		24.8		7359715	0385569	37244	ass',	3738654	791065	37.4644	60.1	24.8	
76	02	ri_cubes	6	0.4	0.0	0.0	TOSI	22	9	14.0	34	15.0	93	5	01	5.0]	97	91	84	581906	8398	829

APPENDIX 3: VERTICAL PEAK GROUND MOTION VARIATIONS BY MAGNITUDE FROM ST1-18_DATABASE_PGM.CSV

Colu mn1	Column 2	Column3	Column4	Column5	Column6	Column7	Column8	Column9	Column10	Column11	Column12	Column13
mag	PGDmax (mm)	PGDmea n(mm)	PGDstde v(mm)	PGDmedia n(mm)	PGVmax(mm/s)	PGVmean(mm/s)	PGVstdev(mm/s)	PGVmedian (mm/s)	PGAmax(mm/s2)	PGAmean(mm/s2)	PGAstdev(mm/s2)	PGAmmedian(mm/s2)
0	5.20E-07	6.17E-08	5.89E-08	4.98E-08	5.18E-04	3.20E-05	5.31E-05	1.96E-05	1.73E-01	1.66E-02	1.90E-02	1.30E-02
0.1	3.86E-07	5.94E-08	4.90E-08	4.62E-08	5.77E-04	3.26E-05	4.57E-05	2.40E-05	1.88E-01	1.64E-02	1.66E-02	1.25E-02
0.2	9.18E-07	8.79E-08	8.41E-08	6.56E-08	6.68E-04	4.20E-05	5.81E-05	2.71E-05	2.43E-01	1.97E-02	2.29E-02	1.34E-02
0.3	1.07E-06	1.02E-07	1.00E-07	7.92E-08	7.34E-04	5.12E-05	6.33E-05	3.37E-05	3.50E-01	2.50E-02	2.72E-02	1.85E-02
0.4	1.66E-06	1.18E-07	1.21E-07	8.61E-08	7.05E-04	6.08E-05	7.34E-05	3.92E-05	5.48E-01	3.08E-02	3.66E-02	2.12E-02
0.5	1.75E-06	1.51E-07	1.48E-07	1.13E-07	6.95E-04	7.25E-05	8.07E-05	4.75E-05	7.14E-01	3.55E-02	4.10E-02	2.50E-02
0.6	3.59E-06	1.84E-07	1.93E-07	1.42E-07	8.77E-04	9.09E-05	1.00E-04	5.88E-05	6.53E-01	4.57E-02	5.03E-02	3.19E-02
0.7	1.14E-05	2.20E-07	3.81E-07	1.63E-07	3.93E-03	1.00E-04	1.34E-04	6.62E-05	1.70E+00	5.04E-02	7.19E-02	3.27E-02
0.8	4.20E-06	3.55E-07	3.97E-07	2.45E-07	2.54E-03	1.61E-04	1.93E-04	1.03E-04	1.91E+00	8.14E-02	1.09E-01	5.26E-02
0.9	1.11E-05	5.23E-07	5.89E-07	4.09E-07	2.88E-03	2.25E-04	2.39E-04	1.56E-04	1.33E+00	1.09E-01	1.17E-01	7.91E-02
1	5.30E-06	4.15E-07	4.20E-07	3.23E-07	1.59E-03	1.87E-04	1.84E-04	1.32E-04	1.06E+00	9.36E-02	9.88E-02	7.05E-02
1.1	5.62E-06	6.28E-07	6.51E-07	4.55E-07	3.07E-03	2.86E-04	3.32E-04	1.88E-04	1.24E+00	1.25E-01	1.33E-01	8.66E-02
1.2	7.20E-06	8.44E-07	8.35E-07	5.90E-07	2.85E-03	3.61E-04	3.98E-04	2.40E-04	4.19E+00	1.80E-01	2.95E-01	1.19E-01
1.3	1.59E-05	1.40E-06	1.49E-06	1.01E-06	4.16E-03	5.56E-04	5.39E-04	3.95E-04	3.23E+00	2.47E-01	2.84E-01	1.75E-01
1.4	1.69E-05	1.35E-06	1.47E-06	1.02E-06	6.07E-03	5.60E-04	6.30E-04	3.77E-04	2.24E+00	2.28E-01	2.46E-01	1.50E-01
1.6	1.23E-05	2.33E-06	1.88E-06	1.87E-06	4.94E-03	9.17E-04	9.24E-04	5.98E-04	2.33E+00	3.99E-01	3.98E-01	2.87E-01
1.7	8.34E-06	1.85E-06	1.14E-06	1.68E-06	3.18E-03	5.57E-04	4.82E-04	4.16E-04	1.70E+00	2.28E-01	2.44E-01	1.52E-01
1.8	1.95E-05	3.09E-06	2.52E-06	2.47E-06	6.45E-03	1.06E-03	1.03E-03	7.12E-04	3.77E+00	3.65E-01	3.99E-01	2.50E-01